

Decoupling of the Kontsevich-Zorich cocycle modulo q and uniform spectral gap

Michael Magee

October 18, 2016

Abstract

We obtain a qualitative extension of Selberg's 3/16 Theorem to the setting of moduli spaces of abelian differentials on genus $g \geq 2$ surfaces. More precisely, under certain conditions we prove that there is a uniform spectral gap for the foliated hyperbolic Laplacian operators associated to congruence covers of a fixed component of a stratum of the moduli space of abelian differentials on a genus g surface. We also state our result in representation-theoretic terms. To prove our result, we build on the work of Avila, Gouëzel and Yoccoz [3] on exponential mixing of the Teichmüller flow as well as the group expansion technology developed in [9, 26, 20]. Our results apply to hyperelliptic components and also extend to arbitrary components on assumption of a conjecture of Zorich that asserts the Zariski-density of the associated Rauzy-Veech group in its ambient symplectic group.

1 Introduction

Let \mathcal{M}_g be the moduli space of abelian differentials on a genus $g \geq 1$ surface S_g . This space is stratified into moduli spaces \mathcal{M}_κ of differentials with prescribed ramification data κ ; see Section 2.1 for background. The space \mathcal{M}_κ may have several components that are understood by work of Kontsevich and Zorich [18]. Let \mathcal{C} be one of these components. An abelian differential on S gives S the structure of a translation surface, let $\mathcal{C}^{(1)}$ denote the differentials in \mathcal{C} with unit area associated translation surface. The spaces $\mathcal{M}_g, \mathcal{M}_\kappa, \mathcal{C}, \mathcal{C}^{(1)}$ have canonical affine orbifold structures arising from period coordinates.

There is an family of finite covering spaces $\mathcal{M}_g(q) \rightarrow \mathcal{M}_g$ parametrized by $q \in \mathbf{N}$ that are of particular number-theoretic interest. Indeed, one can realize $\mathcal{M}_g = \Gamma \backslash \mathcal{X}_g$ where \mathcal{X}_g is a Teichmüller space of differentials and $\Gamma = \Gamma_g$ is the mapping class group of S_g . Then consider the natural homomorphism from Γ_g to the symplectic group $\mathrm{Sp}(H_1(S_g; \mathbf{Z}), \cap)$, where \cap is the intersection form on homology. For each q this induces a map $\Pi_q : \Gamma \rightarrow \mathrm{Sp}(H_1(S_g; \mathbf{Z}/q\mathbf{Z}), \cap)$. We denote by $\Gamma(q)$ the kernel of Π_q , this is a finite index normal subgroup of Γ and hence yields a covering map of orbifolds

$$\mathcal{M}_g(q) := \Gamma(q) \backslash \mathcal{X}_g \rightarrow \mathcal{M}_g. \quad (1.1)$$

As a precursor to the main theorems of this paper, it is natural to ask whether the preimage $\mathcal{C}^{(1)}(q)$ of $\mathcal{C}^{(1)}$ in $\mathcal{M}_g(q)$ under the mapping in (1.1) is connected. This is in fact the case for all q coprime to a fixed modulus $q_0 = q_0(\mathcal{C})$ for the following reasons.

- We introduce in Section 2.2 the *Hodge bundle* $H_1(\mathcal{C}^{(1)})$ over $\mathcal{C}^{(1)}$ that away from orbifold points of $\mathcal{C}^{(1)}$ has fibres modeled on $H_1(S, \mathbf{R})$ and whose structure group is reduced to $\mathrm{Sp}_{2g}(\mathbf{Z})$ by a natural flat Gauss-Manin connection. Then $\mathcal{C}^{(1)}(q)$ is the total space of the associated bundle obtained from the reduction mod q map $\mathrm{Sp}_{2g}(\mathbf{Z}) \rightarrow \mathrm{Sp}_{2g}(\mathbf{Z}/q\mathbf{Z})$.
- The image $\Gamma_{\mathcal{C}}$ of the monodromy representation of the fundamental group of $\mathcal{C}^{(1)}$ associated to the Hodge bundle is known to be Zariski-dense in $\mathrm{Sp}_{2g}(\mathbf{Z})$ by a result of Filip from [11, Corollary 1.3].
- The strong approximation result of Matthews, Vaserstein and Weisfeiler [24] says that if $\Gamma_{\mathcal{C}}$ is Zariski-dense in $\mathrm{Sp}_{2g}(\mathbf{Z})$ then $\Gamma_{\mathcal{C}}$ maps onto $\mathrm{Sp}_{2g}(\mathbf{Z}/q\mathbf{Z})$ for all q coprime to some fixed modulus q_0 and hence that $\mathcal{C}^{(1)}(q)$ is connected for the same q .

There is a well-studied action of $\mathrm{SL}_2(\mathbf{R})$ on $\mathcal{C}^{(1)}$. The restriction of this action to the one parameter diagonal subgroup of $\mathrm{SL}_2(\mathbf{R})$ is called the *Teichmüller flow*, and by work of Masur [23] and Veech [32] there is a unique probability measure $\nu_{\mathcal{C}^{(1)}}$ on $\mathcal{C}^{(1)}$ that is in the Lebesgue class and invariant and ergodic for the Teichmüller flow. This measure is also $\mathrm{SL}_2(\mathbf{R})$ -invariant.

The cocycle over the Teichmüller flow induced by monodromy in the Hodge bundle is called the *Kontsevich-Zorich* (KZ) cocycle and after suitable identifications and lifting takes values in $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_{\mathcal{C}})$ where $\omega_{\mathcal{C}}$ is a symplectic form defined over \mathbf{Z} . The group

$$G_{\mathcal{C}} \subset \mathrm{Sp}(\mathbf{Z}^{2g}, \omega_{\mathcal{C}})$$

generated by the image of the KZ cocycle is called the *Rauzy-Veech group* of \mathcal{C} . Note the distinction between $\Gamma_{\mathcal{C}}$ (all monodromy) and $G_{\mathcal{C}}$ (roughly speaking, monodromy along the Teichmüller flow). The Zariski-density of $G_{\mathcal{C}}$ is a conjecture of Zorich:

Conjecture 1.1 (Zorich [35, Appendix A.3 Conjecture 5]). *For any component \mathcal{C} the group $G_{\mathcal{C}}$ is Zariski-dense in $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_{\mathcal{C}})$.*

Remark 1.2. The Zorich conjecture is related to the perhaps more well known Kontsevich-Zorich conjecture that the Lyapunov exponents of the KZ cocycle w.r.t. $\nu_{\mathcal{C}^{(1)}}$ are semisimple. This is now a theorem due to Avila and Viana [4]. By the comments in [1, Introduction] and reference therein to work of Benoist [6], the Zorich conjecture implies via [4] the Kontsevich-Zorich conjecture but the converse does not hold. The results of this paper can be rephrased in terms of the statistical properties of the KZ cocycle modulo q and therefore can be regarded as in the same vein as the Avila-Viana theorem, but in finite characteristic.

In recent work [1], Avila, Matheus and Yoccoz proved the following strong form of the Zorich conjecture for *hyperelliptic* \mathcal{C} (see Section 2.1 for the definition).

Theorem 1.3 (Avila, Matheus, Yoccoz). *If \mathcal{C} is hyperelliptic the group $G_{\mathcal{C}}$ is finite index in $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_{\mathcal{C}})$.*

There are two hyperelliptic \mathcal{C} for each genus $g \geq 2$. Our main theorems will apply either to hyperelliptic \mathcal{C} where we can apply Theorem 1.3 or to arbitrary \mathcal{C} when we assume the Zorich conjecture.

The main theorem of this paper can be stated in several different ways. The action of $\mathrm{SL}_2(\mathbf{R})$ on $\mathcal{C}^{(1)}$ lifts to an action on $\mathcal{C}^{(1)}(q)$ that preserves the measure $\nu_{\mathcal{C}^{(1)}(q)}$ obtained by

pullback of $\nu_{\mathcal{C}^{(1)}}$. Thus the spaces $\mathrm{SO}(2)\backslash\mathcal{C}^{(1)}(q)$ are foliated by quotients of the hyperbolic plane $\mathrm{SO}(2)\backslash\mathrm{SL}_2(\mathbf{R}) \cong \mathbb{H}$. Using the hyperbolic metric of constant curvature -1 on these leafs, it is possible to define leafwise Laplacian operators¹ Δ_q on each $\mathrm{SO}(2)\backslash\mathcal{C}^{(1)}(q)$. This Δ_q is a nonnegative unbounded operator with dense domain in² $L^2(\mathrm{SO}(2)\backslash\mathcal{C}^{(1)}(q))$. We write $\mathrm{spec}(\Delta_q) \subset [0, \infty)$ for the spectrum of this operator. The first version of our main theorem is the following.

Theorem 1.4 (Uniform spectral gap). *For $g \geq 2$, let \mathcal{C} be a connected component of a stratum of the moduli space of abelian differentials on a genus g surface, and Δ_q the foliated Laplacian associated to $\mathcal{C}^{(1)}(q)$. There is some $q_0 \in \mathbf{N}$ and $\epsilon > 0$ such that*

$$\mathrm{spec}(\Delta_q) \cap [0, \epsilon) = \{0\}$$

for

- all q with $(q, q_0) = 1$ if \mathcal{C} is hyperelliptic, or
- squarefree q with $(q, q_0) = 1$ for any \mathcal{C} , conditional on Conjecture 1.1 for \mathcal{C} .

Another way to state Theorem 1.4 uses Bargmann's classification [5] of the unitary dual of $\mathrm{SL}_2(\mathbf{R})$ into the trivial representation, the principal series, complementary series, discrete series and limits of discrete series. Of particular interest to us are the complementary series Comp^u that are indexed by a parameter $u \in (0, 1)$, see [17, pg. 36] for a precise description of these representations. Let $L_0^2(\mathcal{C}^{(1)}(q))$ be the $\mathrm{SL}_2(\mathbf{R})$ -invariant subspace of functions in $L^2(\mathcal{C}^{(1)}(q))$ that are orthogonal to all lifts of functions in $L^2(\mathcal{C}^{(1)})$ by the map

$$\mathcal{C}^{(1)}(q) \rightarrow \mathcal{C}^{(1)}.$$

We can decompose the representation of $\mathrm{SL}_2(\mathbf{R})$ on $L_0^2(\mathcal{C}^{(1)}(q))$ as a direct integral over the unitary dual of $\mathrm{SL}_2(\mathbf{R})$ with respect to a projection valued measure $\mu_{\mathcal{C}^{(1)}(q)}$. We can then restate Theorem 1.4 as follows.

Theorem 1.5. *Let \mathcal{C} , q_0 and q be as in Theorem 1.4. There is some $\eta > 0$ (independent of q) such that the support of $\mu_{\mathcal{C}^{(1)}(q)}$ is disjoint from those complementary series representations Comp^u with $u \in [1 - \eta, 1)$.*

Theorems 1.4 and 1.5 were proven for $\mathcal{C}^{(1)}$ by Avila, Gouëzel and Yoccoz [3] and we build on their work here. As in [3], we prove Theorem 1.4 by means of the following type of estimate.

Theorem 1.6. *Let \mathcal{C} , q_0 and q be as in Theorem 1.4. The Teichmüller flow on $\mathcal{C}^{(1)}(q)$ has exponential decay of correlations on compactly supported C^1 observables with an error term that depends only polynomially on q .*

¹More discussion on the construction of this operator can be found in [2, Section 3.4], in particular, the foliated Laplacian corresponds to the Casimir operator of $\mathrm{SL}_2(\mathbf{R})$ acting on the $\mathrm{SO}(2)$ -invariant functions in $L^2(\mathcal{C}^{(1)}(q))$.

²When we discuss L^2 spaces we always mean with respect to the relevant $\nu_{\mathcal{C}^{(1)}(q)}$ or $\nu_{\mathcal{C}^{(1)}}$, or in the case of $\mathrm{SO}(2)\backslash\mathcal{C}^{(1)}(q)$, the measure induced by the $\mathrm{SO}(2)$ -invariant $\nu_{\mathcal{C}^{(1)}(q)}$.

It should be possible to extend Theorem 1.6 to Hölder observables of suitable decay as in [3], however we do not need this for Theorem 1.4 and other potential applications, so we stick to a more restricted class to simplify the paper. We give a precise formulation of Theorem 1.6 in Theorem 3.5. Theorem 3.5 implies Theorem 1.5 by the ‘reverse Ratner estimates’ given in [3, Appendix B], and the connection between Theorems 1.4 and 1.5 is discussed in detail in [2, Section 3.4].

Selberg’s celebrated 3/16 Theorem [30] gives a quantitative form of Theorem 1.4 for the Laplace-Beltrami operators Δ_q on the finite volume arithmetic Riemann surfaces

$$X(q) := \Gamma(q) \backslash \mathbb{H},$$

where $\Gamma(q)$ is the kernel of reduction modulo q on $\mathrm{SL}_2(\mathbf{Z})$. Indeed, Selberg proved that in this setting

$$\mathrm{spec}(\Delta_q) \cap [0, 3/16] = \{0\}, \quad (1.2)$$

and furthermore conjectured that the same holds with 3/16 replaced by 1/4. Selberg’s conjecture remains a fundamental open problem of automorphic forms; the current best bound is due to Kim and Sarnak [16, Appendix 2] that allows 975/4096 to replace 3/16 in (1.2). The statement of Theorem 1.4 for $g = 1$ and arbitrary q is exactly Selberg’s Theorem with ϵ replacing 3/16 since

1. When $g = 1$, the only stratum \mathcal{M}_\emptyset is connected and the corresponding action of $\mathrm{SL}_2(\mathbf{R})$ on $\mathcal{C}^{(1)}$ is exactly the left multiplication action of $\mathrm{SL}_2(\mathbf{R})$ on $\mathrm{SL}_2(\mathbf{R})/\mathrm{SL}_2(\mathbf{Z})$.
2. In this case, the foliated Laplacian Δ_q is the Laplace-Beltrami operator on the modular curve $X(q)$ that appears in Selberg’s theorem.

So Theorem 1.4 can be regarded as a qualitative extension of Selberg’s Theorem to higher genus.

Bourgain, Gamburd and Sarnak [9] obtained a similar qualitative generalization of Selberg’s 3/16 Theorem to the setting where Γ is a finitely generated, infinite index and Zariski-dense (*thin*) subgroup of $\mathrm{SL}_2(\mathbf{Z})$, $\Gamma(q)$ is the kernel of reduction modulo q on Γ , and q is squarefree and coprime to some modulus $q_0 = q_0(\Gamma)$. They also required that the Hausdorff dimension $\delta(\Gamma)$ of the limit set of an orbit $\Gamma.o$ in $\delta\mathbb{H}$ satisfies $\delta > 1/2$.

Further work of Oh and Winter [26] then treated congruence covers of thin $\Gamma \subset \mathrm{SL}_2(\mathbf{Z})$ with $0 < \delta \leq 1/2$, but with the same restriction that q is squarefree coprime to a finite modulus. In fact, Oh and Winter work primarily with the analog of Theorem 1.6, and it is their approach we follow here, building on their methods. An argument that allows the restriction that q is squarefree to be removed was given by Bourgain, Kontorovich and Magee in [20, Appendix]. The current paper brings these ideas into the setting of Teichmüller dynamics.

A key ingredient in the proof of Theorem 1.4 in the case when \mathcal{C} is hyperelliptic is Kazhdan’s property (T) from [14]. A locally compact group G has property (T) if its trivial representation is isolated in the Fell topology on the unitary dual of G . We make use of the following result of Kazhdan [14].

Theorem 1.7 (Kazhdan). *The group $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_{\mathcal{C}})$, and also any finite index subgroup, has Kazhdan’s property (T).*

See Lemma 4.9 for the use of Theorem 1.7 in this paper. Whenever we are only assuming $G_{\mathcal{C}}$ is Zariski-dense but not finite index, we do not have access to property (T)³ so rely instead on the following result of Salehi Golsefidy and Varjú [12].

Theorem 1.8 (Salehi Golsefidy, Varjú). *There is finite q_0 such that the Cayley graphs of $\mathrm{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_{\mathcal{C}})$ with respect to the projection of any fixed set of generators of $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_{\mathcal{C}})$ form a family of uniform expander graphs as q ranges over squarefree positive integers coprime to q_0 .*

The statement of Theorem 1.8 is translated into a spectral gap result in Lemma 4.9. The decoupling arguments of [20, Appendix] that we use here also depend on lower bounds for the dimensions of irreducible representations of $\mathrm{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_{\mathcal{C}})$ that do not arise from representations of $\mathrm{Sp}((\mathbf{Z}/q_1\mathbf{Z})^{2g}, \omega_{\mathcal{C}})$ with $q_1|q$. This is a version of the *quasirandomness* property⁴ of a group that takes into account the level structure of the family of groups $\mathrm{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_{\mathcal{C}})$.

Proposition 1.9 (Quasirandomness estimates). *There is $C > 0$ and $D > 0$ such that any irreducible representation of $\mathrm{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_{\mathcal{C}})$ that does not factor through*

$$\mathrm{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_{\mathcal{C}}) \rightarrow \mathrm{Sp}((\mathbf{Z}/q_1\mathbf{Z})^{2g}, \omega_{\mathcal{C}})$$

for some $q_1|q$ has dimension $\geq Cq^D$.

We explain how Proposition 1.9 is proved in Section 5, following an argument of Kelmer and Silberman [15].

1.1 Acknowledgements

We would like to thank Alex Eskin, Ilya Gekhtman, Babak Modami, Hee Oh and Giulio Tiozzo for helpful discussions related to this work.

2 Background

2.1 Abelian differentials and translation surfaces

Let $g \geq 1$ and let $S = S_g$ be a fixed topological surface of genus g . An *abelian differential* on S is a pair (\mathcal{J}, ω) where \mathcal{J} is a complex structure on S and ω is a holomorphic one form with respect to that structure. As is well known, an abelian differential ω on S gives S the structure of a translation surface with conical singularities at the zeros of ω . The homeomorphisms of S , $\mathrm{Homeo}(S)$, act on the abelian differentials by pushforward of the complex structure and differential.

³In fact, we think it is an open problem whether there are infinite index Zariski-dense subgroups of $\mathrm{Sp}_{2g}(\mathbf{Z})$ with property (T) for $g \geq 2$.

⁴The following notion arises in work of Gowers [13, Theorem 4.5]: that a finite group G should be regarded as quasirandom relative to an ambient parameter C if the dimension of any nontrivial irreducible representation of G has dimension $\geq C$. Prior to this formal notion, the concept had been used in the construction of Ramanujan graphs by Lubotzky, Phillips and Sarnak [19], the work of Sarnak and Xue on multiplicities of automorphic representations [28], and the construction of uniformly expanding Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$ by Bourgain and Gamburd [8].

We define the *Teichmüller space* \mathcal{X}_g of abelian differentials of genus g to be the quotient of the space of abelian differentials on S by homeomorphisms of S isotopic to the identity, that we denote $\text{Homeo}_0(S)$. The *moduli space* \mathcal{M}_g of abelian differentials is defined to be the quotient of the abelian differentials on S by $\text{Homeo}(S)$. Thus there is an action of the mapping class group

$$\Gamma = \Gamma_g := \text{Homeo}(S)/\text{Homeo}_0(S)$$

on \mathcal{X}_g such that

$$\Gamma_g \backslash \mathcal{X}_g = \mathcal{M}_g.$$

The moduli space \mathcal{M}_g is stratified according to the orders of the zeros of the abelian differentials: for $\kappa = (\kappa_1, \dots, \kappa_s)$ with $\sum \kappa_i = 2g - 2$ we let \mathcal{X}_κ (resp. \mathcal{M}_κ) denote the Teichmüller space (resp. moduli space) of abelian differentials with s zeros of orders $\kappa_1, \dots, \kappa_s$ (for some ordering of the zeros). Then \mathcal{M}_κ is often referred to as a *stratum* of the moduli space \mathcal{M}_g . The connected components of these strata have been classified by Kontsevich and Zorich [18].

The classification of Kontsevich and Zorich involves a special type of component that is called *hyperelliptic* [18, Definition 2]. A hyperelliptic complex structure on S is one that has an involution such that the quotient of (S, \mathcal{J}) by the involution is \mathbb{CP}^1 . A hyperelliptic component is one consisting of (equivalence classes of) pairs (\mathcal{J}, ω) where \mathcal{J} is a hyperelliptic structure. These occur for $g \geq 2$ either as a component of \mathcal{M}_κ with $\kappa = (2g - 2)$ or a component of \mathcal{M}_κ with $\kappa = (g - 1, g - 1)$. In both cases the zero(s) of the abelian differential are symmetric w.r.t. the hyperelliptic involution.

The Teichmüller space \mathcal{X}_κ has a complex manifold structure that comes from period coordinates as in [3, Section 2.2.1]. Any connected component \mathcal{C} of \mathcal{M}_κ inherits the structure of a complex affine orbifold. The subspace $\mathcal{X}_\kappa^{(1)} \subset \mathcal{X}_\kappa$ consisting of unit area translation surfaces is a real hypersurface, accordingly the set of moduli $\mathcal{C}^{(1)} \subset \mathcal{C}$ of translation surfaces with unit area is a real affine orbifold. Returning to the discussion of the Introduction, for each q the lift of $\mathcal{C}^{(1)}(q)$ to \mathcal{X}_κ is a submanifold of $\mathcal{X}_\kappa^{(1)}$.

Recall that a Finsler manifold is a smooth manifold together with a continuous assignment of norm on each tangent fibre. The norm is called a Finsler metric. As described in [3, Section 2.2.2] there is a natural Finsler metric on $\mathcal{X}_\kappa^{(1)}$ making it into a Finsler manifold.

2.2 The Hodge bundle

Recall that \mathcal{X}_κ is the Teichmüller space of abelian differentials on S with ramification data κ . The Hodge bundle is defined to be the fibred product

$$H_1(\mathcal{M}_\kappa) := \Gamma \backslash (\mathcal{X}_\kappa \times H_1(S, \mathbf{R})) \rightarrow \Gamma \backslash \mathcal{X}_\kappa = \mathcal{M}_\kappa$$

where the mapping class group Γ acts diagonally. Let \mathcal{M}_κ^0 be the complement of the orbifold points in \mathcal{M}_κ . The Hodge bundle restricts to a vector bundle $H_1(\mathcal{M}_\kappa^0)$ over \mathcal{M}_κ^0 . At any orbifold point $[(\mathcal{J}, \omega)]$ of \mathcal{M}_κ the fibre degenerates to $\text{Aut}(\mathcal{J}, \omega) \backslash H_1(S, \mathbf{R})$. Note that by Hurwitz's automorphisms theorem, $\text{Aut}(\mathcal{J}, \omega)$ is a finite group.

The total space of the Hodge bundle contains as a discrete subset the lattice bundle

$$\Gamma \backslash (\mathcal{X}_\kappa \times H_1(S, \mathbf{Z})).$$

Then one may specify the *Gauss-Manin* connection on the Hodge bundle by the requirement that lattice valued continuous sections be parallel. This gives a flat vector bundle connection on $H_1(\mathcal{M}_\kappa^0)$ that extends to a flat connection on $H_1(\mathcal{M}_\kappa)$ in the following sense. A section of $H_1(\mathcal{M}_\kappa)$ can be viewed as a function $\sigma : \mathcal{X}_\kappa \rightarrow H_1(S, \mathbf{R})$ that transforms according to

$$\sigma(\gamma.x) = \gamma_*\sigma(x), \quad \gamma \in \Gamma.$$

Then a local section is parallel by definition if it takes values in $H_1(S, \mathbf{Z})$ and this specifies the connection on general sections.

The action of Γ on $H_1(S, \mathbf{Z})$ lies in the integral symplectic group $\mathrm{Sp}(H_1(S, \mathbf{Z}), \cap)$ where \cap is the (symplectic) intersection form on integral homology. Therefore for any unitary representation (ρ, V) of $\mathrm{Sp}(H_1(S, \mathbf{Z}), \cap)$ we obtain an *associated orbifold vector bundle*⁵ $H_1(\mathcal{M}_\kappa; \rho)$. The total space of this bundle is

$$H_1(\mathcal{M}_\kappa; \rho) = \Gamma \backslash (\mathcal{X}_\kappa \times V) \tag{2.1}$$

where the action of Γ on $\mathcal{X}_\kappa \times V$ is given by $\gamma.(\omega, v) = (\gamma.\omega, \rho(\gamma_*)v)$, where $\gamma_* \in \mathrm{Sp}(H_1(S, \mathbf{Z}), \cap)$ is the map induced by γ on homology. This bundle also has a flat connection, in the same sense as before, coming from the fibred product structure in (2.1).

Of course, for any connected component \mathcal{C} of the stratum \mathcal{M}_κ we may restrict $H_1(\mathcal{M}_\kappa)$ or $H_1(\mathcal{M}_\kappa; \rho)$ to \mathcal{C} or even $\mathcal{C}^{(1)}$. We denote by $H_1(\mathcal{C}; \rho)$ or $H_1(\mathcal{C}^{(1)}; \rho)$ the obtained orbifold vector bundles.

For a lot of the rest of the paper we deal with abstract unitary ρ but in reality we are interested in the following specific examples. Recall the maps $\Pi_q : \Gamma \rightarrow \mathrm{Sp}(H_1(S_g; \mathbf{Z}/q\mathbf{Z}), \cap)$ from the Introduction and let $\Gamma_q := \Gamma / \ker \Pi_q$. *Throughout the rest of the paper when we reference q we assume that strong approximation holds at q and hence*

$$\Gamma_q \cong \mathrm{Sp}(H_1(S_g; \mathbf{Z}/q\mathbf{Z}), \cap).$$

Let $\ell_0^2(\Gamma_q)$ be the subspace of functions in $\ell^2(\Gamma_q)$ that are orthogonal to constant functions with respect to the ℓ^2 inner product. This gives a subrepresentation $(\rho_q, \ell_0^2(\Gamma_q))$ of the action of Γ on $\ell^2(\Gamma_q)$ by reduction mod q and then left translation⁶.

We will also consider the subspace of $\ell_0^2(\Gamma_q)$ consisting of functions that are orthogonal to all functions lifted from $\Gamma_{q'}$ with $q'|q$ via the natural mapping of reduction modulo q'

$$\Gamma_q \rightarrow \Gamma_{q'}.$$

We denote by $\ell_{\mathrm{new}}^2(\Gamma_q)$ this *new subspace* of functions. This gives a subrepresentation $(\rho_q^{\mathrm{new}}, \ell_{\mathrm{new}}^2(\Gamma_q))$ of $(\rho_q, \ell_0^2(\Gamma_q))$.

2.3 The Teichmüller flow on moduli space

There is a postcomposition action of $\mathrm{SL}_2(\mathbf{R})$ on the space of abelian differentials on S as follows. For $h \in \mathrm{SL}_2(\mathbf{R})$ we define

⁵By *orbifold vector bundle* we mean that the fibres are vector spaces of constant rank away from the orbifold points of the base space, where the fibres degenerate only to a quotient of a vector space by a finite group.

⁶In other words, the inflation of the left regular representation of Γ_q to Γ .

$$h.(\mathcal{J}, \omega) = (\mathcal{J}_h, \omega_h)$$

where

$$\omega_h = h \begin{pmatrix} \Re(w) \\ \Im(w) \end{pmatrix}$$

and \mathcal{J}_h is the unique complex structure on S that makes ω_h holomorphic. As this action commutes with any homeomorphism of S , it descends to both Teichmüller space \mathcal{X}_g and moduli space \mathcal{M}_g , moreover the action preserves any connected component of any stratum \mathcal{M}_κ . The *Teichmüller geodesic flow* on any of these objects is the restriction of the $\mathrm{SL}_2(\mathbf{R})$ action to the diagonal subgroup:

$$\mathcal{T}_t(\mathcal{J}, \omega) := \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}.(\mathcal{J}, \omega).$$

The Teichmüller flow preserves each connected component \mathcal{C} of a stratum and moreover preserves the unit area surfaces $\mathcal{C}^{(1)}$. By results of Masur [23] and Veech [32] there is a unique probability measure $\nu_{\mathcal{C}^{(1)}}$ that is invariant and ergodic for the Teichmüller flow on $\mathcal{C}^{(1)}$. This measure is in the Lebesgue class with respect to period coordinates on $\mathcal{C}^{(1)}$.

Since in Section 2.2 we specified a connection on each of $H_1(\mathcal{M}_\kappa)$, $H_1(\mathcal{M}_\kappa; \rho)$ the Teichmüller flow acts on sections of each of these bundles by pullback along parallel transport. For example, viewing a section of $H_1(\mathcal{M}_\kappa; \rho)$ as a V -valued function σ on \mathcal{X}_κ satisfying $\sigma(\gamma.x) = \rho(\gamma)\sigma(x)$ for each $\gamma \in \Gamma$, we have the following defining equation for \mathcal{T}_t^* :

$$[\mathcal{T}_t^* \sigma](\mathcal{J}, \omega) := \sigma(\mathcal{T}_t(\mathcal{J}, \omega)). \quad (2.2)$$

This action also restricts to an action on sections of $H_1(\mathcal{C}; \rho)$ and $H_1(\mathcal{C}^{(1)}; \rho)$. We now explain the relationship between sections of $H_1(\mathcal{C}^{(1)}, \rho_q)$ and functions on $\mathcal{C}^{(1)}(q)$. Recall that $L_0^2(\mathcal{C}^{(1)}(q))$ is the subspace of functions in $L^2(\mathcal{C}^{(1)}(q))$ orthogonal to lifts from $L^2(\mathcal{C}^{(1)})$, w.r.t. the measure $\nu_{\mathcal{C}^{(1)}(q)}$. Let $L^2(H_1(\mathcal{C}^{(1)}, \rho_q))$ denote the L^2 sections of $H_1(\mathcal{C}^{(1)}, \rho_q)$ w.r.t. the natural Hermitian fibre metric and measure $\nu_{\mathcal{C}^{(1)}}$. We say that a function f on $\mathcal{C}^{(1)}(q)$ or a section σ of $H_1(\mathcal{C}^{(1)}, \rho)$ is C^1 if its lift to $\tilde{f} : \mathcal{X}_\kappa^{(1)} \rightarrow \mathbf{C}$ (resp. $\tilde{\sigma} : \mathcal{X}_\kappa^{(1)} \rightarrow V$) is C^1 (bounded with bounded derivative⁷) w.r.t. the Finsler manifold structure on $\mathcal{X}_\kappa^{(1)}$. Define $\|f\|_{C^1} = \|\tilde{f}\|_\infty + \|D\tilde{f}\|_\infty$ and similarly $\|\sigma\|_{C^1}$. Write $C^1(\mathcal{C}^{(1)}(q))$ for the C^1 complex valued functions on $\mathcal{C}^{(1)}(q)$ and $C^1(H_1(\mathcal{C}^{(1)}, \rho))$ for the C^1 sections of $H_1(\mathcal{C}^{(1)}, \rho)$. These are Banach spaces w.r.t the respective C^1 norms.

Lemma 2.1. *We have the following correspondences*

1. *For each q there is a natural linear isometry*

$$\Phi_q : L_0^2(\mathcal{C}^{(1)}(q)) \rightarrow L^2(H_1(\mathcal{C}^{(1)}, \rho_q)).$$

2. *The map Φ_q intertwines the maps \mathcal{T}_t^* defined by pullback on $L_0^2(\mathcal{C}^{(1)}(q))$ and by (2.2) on $L^2(H_1(\mathcal{C}^{(1)}, \rho_q))$.*

⁷In case of V -valued F on a Finsler manifold X with V a Hilbert space, to define the norm of the derivative we view the derivative at $x \in X$ as a map $DF_x : T_x X \rightarrow T_{F(x)} V \cong V$ then use the operator norm w.r.t. the Finsler metric at x and the Hilbert space norm on V .

3. The restriction

$$\Phi_q : C^1(\mathcal{C}^{(1)}(q)) \cap L_0^2(\mathcal{C}^{(1)}(q)) \rightarrow C^1(H_1(\mathcal{C}^{(1)}, \rho_q))$$

preserves C^1 norms.

2.4 Combinatorial data and Rauzy classes

Now we begin an account of the dynamics of the Teichmüller flow, viewed through the lens of Veech's zippered rectangles construction. We draw in the following sections from the sources [3], [33] that both build on work of Marmi, Moussa and Yoccoz [22].

The relevant combinatorial objects are as follows. Let \mathcal{A} denote a finite alphabet with $|\mathcal{A}| = d$. Eventually, \mathcal{A} will be chosen depending on g, κ and the component \mathcal{C} . We let $\mathfrak{S}(\mathcal{A})$ denote the set of pairs

$$(\pi_t, \pi_b)$$

where each $\pi_\epsilon : \mathcal{A} \rightarrow \{1, \dots, d\}$. Henceforth, ϵ will index one of the symbols t, b ('top' or 'bottom'). As in [3] it is convenient to visualize (π_t, π_b) as a pair of rows each of which contains the elements of \mathcal{A} in some order, where the top corresponds to π_t and the bottom to π_b . We say (π_t, π_b) is *irreducible* if there is no $d' < d$ such that the set of the first d' elements of the top row is the same as the first d' elements of the bottom. Let $\mathfrak{S}^0(\mathcal{A}) \subset \mathfrak{S}(\mathcal{A})$ denote the irreducible combinatorial data.

We now define 'top' and 'bottom' operations on $\mathfrak{S}^0(\mathcal{A})$. For the next paragraph, let α and β denote the last elements of the top and bottom rows of $\pi \in \mathfrak{S}^0(\mathcal{A})$ respectively. The top operation on π modifies the bottom row by moving the occurrence of β to the immediate right of the occurrence of α . The bottom operation modifies the top row by moving α to the right of β . As in [3] we say that the last element of the unchanged row is the *winner* and the last element of the row of π that is to be changed the *loser*.

By adding directed 'top' and 'bottom' labelled edges according to these operations we obtain an edge-labeled directed graph on the vertex set of irreducible combinatorial data $\mathfrak{S}^0(\mathcal{A})$. Each vertex has exactly one incoming top (resp. bottom) and one outgoing top (resp. bottom) edge. A *Rauzy diagram* is a connected component of this graph and a *Rauzy class* is the vertex set of a Rauzy diagram.

2.5 Suspension data and zippered rectangles

Let be a Rauzy class. For each $\pi \in \mathfrak{R}$ we form a cell

$$X_\pi = \{\pi\} \times \mathbf{R}_+^A \times \mathcal{K}_\pi$$

where

$$\mathcal{K}_\pi = \left\{ \tau \in \mathbf{R}^A : \sum_{\pi_t(\xi) \leq k} \tau_\xi > 0, \sum_{\pi_b(\xi) \leq k} \tau_\xi < 0 \text{ for all } 1 \leq k \leq d-1. \right\}$$

The set \mathcal{K}_π is an open convex cone. Let $X_{\mathfrak{R}} = \bigcup_{\pi \in \mathfrak{R}} X_\pi$. We may drop the dependence on \mathfrak{R} since we usually view it as fixed. We associate to each $\pi \in \mathfrak{R}$ a linear map $\Omega_\pi : \mathbf{R}^A \rightarrow \mathbf{R}^A$ given by

$$[\Omega_\pi]_{\alpha,\beta} = \begin{cases} +1 & \text{if } \pi_t(\alpha) > \pi_t(\beta), \pi_b(\alpha) < \pi_b(\beta), \\ -1 & \text{if } \pi_t(\alpha) < \pi_t(\beta), \pi_b(\alpha) > \pi_b(\beta), \\ 0 & \text{else.} \end{cases}$$

There is a construction due to Veech [32] that builds a point in the moduli space of translation surfaces from suspension data. This mapping is called the *zippered rectangles* construction that we denote by

$$\text{zip} : X_{\mathfrak{R}} \rightarrow \mathcal{M}_\kappa, \quad \kappa = \kappa(\mathfrak{R}).$$

The explicit details of this construction are clearly described in lecture notes of Viana [33, Chapter 2]. In the current paper it will be better to simply work with the properties of the map zip that we give below.

Theorem 2.2 (Veech [32]). *For any connected component \mathcal{C} of the stratum \mathcal{M}_κ there is a Rauzy class $\mathfrak{R} = \mathfrak{R}(\mathcal{C})$ such that $\text{zip}(X_{\mathfrak{R}}) \subset \mathcal{C}$ and $\text{zip}(X_{\mathfrak{R}})$ has full measure w.r.t $\nu_{\mathcal{C}(1)}$.*

There is a natural identification

$$\mathbf{R}^A / \ker \Omega_\pi \cong H_1(\text{zip}(\pi, \lambda, \tau), \mathbf{R}) \quad (2.3)$$

for each $(\pi, \lambda, \tau) \in X_\pi$. This descends to an isomorphism of integral symplectic lattices

$$(\mathbf{Z}^A / \ker(\Omega_\pi|_{\mathbf{Z}^A}), \omega_\pi) \cong (H_1(S, \mathbf{Z}), \cap). \quad (2.4)$$

Therefore the pull back of the Hodge bundle to X_π via zip is naturally trivialized:

$$[\text{zip}^* H_1(\mathcal{M}_\kappa)]|_{X_\pi} \cong X_\pi \times \mathbf{R}^A / \ker \Omega_\pi. \quad (2.5)$$

For a detailed discussion of this map see Viana [33, Section 2.9]. The bilinear form

$$(v_1, v_2) \mapsto \langle \lambda, -\Omega_\pi \tau \rangle$$

descends to a nondegenerate symplectic form ω_π on $\mathbf{R}^A / \ker \Omega_\pi$. Under the identification (2.3), the form ω_π is precisely the intersection form on homology. We also note here that the area of $\text{zip}(\pi, \lambda, \tau)$ is given by

$$\text{area}(\text{zip}(\pi, \lambda, \tau)) = \langle \lambda, -\Omega_\pi \tau \rangle. \quad (2.6)$$

2.6 The Rauzy induction map

Given π , let α be the last element of the top row of π and β the last element of the bottom row. Say that a pair (π, λ) has type *top* if $\lambda_\alpha > \lambda_\beta$. Say it has type *bottom* if $\lambda_\beta < \lambda_\alpha$. This splits each cell into two pieces of the form

$$X_{\pi,\epsilon} = \{ (\pi, \lambda, \tau) \in \{\pi\} \times \mathbf{R}_+^A \times \mathcal{K}_\pi : (\pi, \lambda) \text{ of type } \epsilon \}, \quad \epsilon \in \{t, b\}$$

together with a hyperplane. We also introduce $Y_{\pi,\epsilon} = \{ (\pi, \lambda) \in \{\pi\} \times \mathbf{R}_+^A \text{ of type } \epsilon \}$, so that

$$X_{\pi,\epsilon} = Y_{\pi,\epsilon} \times \mathcal{K}_\pi.$$

We now give an assignment of a linear map $\Theta_{\pi,\epsilon} : \mathbf{R}^A \rightarrow \mathbf{R}^A$ to each pair (π, ϵ) . This is given by [33, (1.9),(1.10)]

$$[\Theta_{\pi,\epsilon}]_{\alpha,\beta} := \begin{cases} 1 & \text{if } \alpha=\beta \\ 1 & \text{if } \alpha \text{ loses and } \beta \text{ wins in type } \epsilon \text{ move at } \pi \\ 0 & \text{else.} \end{cases} \quad (2.7)$$

If π' is obtained from π by a type ϵ move then the map⁸ $\Theta_{\pi,\epsilon}^*$ maps $Y_{\pi'} := \{\pi'\} \times \mathbf{R}_+^A$ homeomorphically to $Y_{\pi,\epsilon}$. Furthermore $(\Theta_{\pi,\epsilon}^*)^{-1}$ maps \mathcal{K}_π injectively into $\mathcal{K}_{\pi'}$ [33, Lemma 2.13]. We also have the intertwining relation

$$\Theta_{\pi,\epsilon} \Omega_\pi \Theta_{\pi,\epsilon}^* = \Omega_{\pi'}. \quad (2.8)$$

The *Rauzy induction map* on suspension data is given by

$$\hat{Q}(\pi, \lambda, \tau) := (\pi', (\Theta_{\pi,\epsilon}^*)^{-1}\lambda, (\Theta_{\pi,\epsilon}^*)^{-1}\tau)$$

when $(\pi, \lambda, \tau) \in X_{\pi,\epsilon}$; here again π' is obtained from π by an operation of type ϵ . Using the same notation, notice that \hat{Q} is a skew extension of the map⁹

$$Q(\pi, \lambda) = (\pi', (\Theta_{\pi,\epsilon}^*)^{-1}\lambda).$$

The equation (2.8) together with the area formula (2.6) shows that \hat{Q} preserves the area of the associated zippered rectangles. Moreover, the zippered rectangles associated to (π, λ, τ) define the same point in \mathcal{M}_κ as the zippered rectangles associated to $\hat{Q}(\pi, \lambda, \tau)$, that is,

$$\text{zip} \circ \hat{Q} = \text{zip}.$$

See for example Viana [33, Section 2.8] for a clear explanation of this fact.

We now define cylinders for the Rauzy induction map. Let γ be a path in the Rauzy diagram associated to the class \mathfrak{R} . Throughout the rest of the paper, we consider oriented paths that follow the given direction of the edges¹⁰. Suppose that γ traverses vertices $\pi(0), \pi(1), \dots, \pi(N)$ in that order. Then define

$$X_\gamma := X_{\pi(0)} \cap \hat{Q}^{-1}(X_{\pi(1)}) \cap \hat{Q}^{-2}(X_{\pi(2)}) \cap \dots \cap \hat{Q}^{-N}(X_{\pi(N)}).$$

Notice that $X_{\pi,\epsilon}$ is the same as X_γ where γ is the outgoing type ϵ arrow from π . We then define Θ_γ in terms of the $\Theta_{\pi,\epsilon}$ by stating that for $(\pi, \lambda, \tau) \in X_\gamma$ we have

$$\hat{Q}^N(\pi(0), \lambda, \tau) = (\pi(N), (\Theta_\gamma^*)^{-1}\lambda, (\Theta_\gamma^*)^{-1}\tau).$$

We define $Y_\gamma = Y_{\pi(0)} \cap \dots \cap Q^{-N}(Y_{\pi(N)})$ the analogous cylinder for Q . If γ begins at π then we define the subcone of \mathcal{K}_π

$$\mathcal{K}_\gamma := (\Theta_\gamma^*)^{-1}\mathcal{K}_\pi.$$

⁸Here and henceforth a $*$ denotes a transpose with respect to the standard basis of \mathbf{R}^A .

⁹As a comment for the initiated, the map Q is the Rauzy induction map on Interval Exchange Transformations. See [27] for Rauzy's original analysis of this map.

¹⁰While it is not immediately obvious, the equivalence classes induced by identifying end points of oriented paths coincide with the Rauzy classes [33, Lemma 1.23].

2.7 The Rauzy-Veech group

Observe that $\Theta_{\pi,\epsilon}^*$ induces a map $\mathbf{Z}^A/\ker \Omega_{\pi'} \rightarrow \mathbf{Z}^A/\ker \Omega_{\pi}$ in light of (2.8) and the fact that $\Theta_{\pi,\epsilon}$ is integral from (2.7). These facts are discussed by Viana in [33, Section 2.8]. As a consequence, (2.8) implies that if γ begins and ends at π , Θ_{γ}^* induces a symplectic endomorphism of $(\mathbf{Z}^A/\ker(\Omega_{\pi}|_{\mathbf{Z}^A}), \omega_{\pi}) \cong^{(2.4)} (H_1(S, \mathbf{Z}), \cap)$. In fact it is easy to check from (2.7) that Θ_{γ}^* is an automorphism. We therefore view each

$$\Theta_{\gamma}^* \in \mathrm{Sp}(\mathbf{Z}^{2g}, \omega_{\pi}).$$

For each $\pi \in \mathfrak{R}$ let G_{π} be the subgroup of $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_{\pi})$ generated by the Θ_{γ}^* obtained as γ ranges over loops in \mathfrak{R} beginning and ending at π . This group G_{π} is called the *Rauzy-Veech group* at π . The $G_{\mathcal{C}}$ discussed in the Introduction is G_{π} for some choice of $\pi \in \mathfrak{R}(\mathcal{C})$. We may now restate Conjecture 1.1 in the following more precise way.

Conjecture 2.3 (Zorich [35, Appendix A.3 Conjecture 5]). *For any Rauzy class \mathfrak{R} and $\pi \in \mathfrak{R}$ the group G_{π} is Zariski-dense in $\mathrm{Sp}(\mathbf{C}^{2g}, \omega_{\pi})$.*

Notice that the conjugacy class of G_{π} only depends on the Rauzy class of π , and therefore the Zariski-density statement in Conjecture 2.3 is a feature only of the Rauzy class of π and hence of the component \mathcal{C} .

2.8 Relationship to the Hodge bundle

Let \mathcal{C} be a connected component of \mathcal{M}_{κ} and let σ be a section of the Hodge bundle $H_1(\mathcal{M}_{\kappa})|_{\mathcal{C}}$. The pullback of σ to any X_{π} under the zippered rectangles map can be naturally viewed as a

$$\mathbf{R}^A/\ker \Omega_{\pi}$$

valued function $\tilde{\sigma}$ via the identifications (2.3) and (2.5). Since Rauzy induction does not change the modulus of zippered rectangles, the fibre of $\mathrm{zip}^* H_1(\mathcal{M}_{\kappa})$ at (π, λ, τ) should be identified with the fibre at $R(\pi, \lambda, \tau)$. In fact, the identification involves the previously defined map Θ_{γ} and requires for $(\pi, \lambda, \gamma) \in X_{\pi,\epsilon}$ that if π' is the result of applying a type ϵ move to π then

$$\tilde{\sigma}(\pi, \lambda, \tau) = \Theta_{\pi,\epsilon}^* \tilde{\sigma}(\hat{Q}(\pi, \lambda, \tau)). \quad (2.9)$$

The iterated form of the compatibility equation (2.9) that we will use is the following. If γ is a path of N edges in a Rauzy diagram that begins and ends at π , then for $(\pi, \lambda, \tau) \in X_{\gamma}$ we have

$$\tilde{\sigma}(\pi, \lambda, \tau) = \Theta_{\gamma}^* \tilde{\sigma}(\hat{Q}^N(\pi, \lambda, \tau)).$$

This is an important point of this paper as it describes the equivariance properties of sections of the Hodge bundle in the suspension model. We now extend this formula to the setting of associated orbifold vector bundles $H_1(\mathcal{C}; \rho)$. After fixing π , using the isomorphism (2.4) and recalling the standing strong approximation assumption we identify

$$\Gamma_q \cong \mathrm{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_{\pi})$$

so we may view ρ_q and ρ_q^{new} as representations of $\text{Sp}(\mathbf{Z}^{2g}, \omega_\pi)$ that are submodules of $\ell^2(\text{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_\pi))$. More generally, using (2.4) we may pull back any unitary representation (ρ, V) of $\text{Sp}(H_1(S; \mathbf{Z}), \cap)$ to a representation of $\text{Sp}(\mathbf{Z}^{2g}, \omega_\pi)$ that we also call ρ .

We may now argue by analogy with the Hodge bundle that if σ is any section of the associated bundle $H_1(\mathcal{C}; \rho)$ then the pull back $\tilde{\sigma}$ of this section to a V -valued function on X_π satisfies

$$\tilde{\sigma}(\pi, \lambda, \tau) = \rho(\Theta_\gamma^*) \tilde{\sigma}(\hat{Q}^N(\pi, \lambda, \tau)), \quad (\pi, \lambda, \tau) \in X_\gamma \quad (2.10)$$

for each path γ in \mathfrak{R} of length N beginning and ending at π .

2.9 A fundamental domain

There is a nice fundamental domain for Rauzy induction on X described in [3, pg. 159]. We let $\mathcal{F} = \mathcal{F}_\mathfrak{R}$ denote the set of (π, λ, τ) such that either

1. $\hat{Q}(\pi, \lambda, \tau) = (\pi', \lambda', \tau')$ is defined and $\|\lambda'\| < 1 \leq \|\lambda\|$
2. $\hat{Q}(\pi, \lambda, \tau)$ is not defined and $1 \leq \|\lambda\|$
3. $\hat{Q}^{-1}(\pi, \lambda, \tau)$ is not defined and $\|\lambda\| < 1$.

The norm we use is $\|\lambda\| := \sum_{\alpha \in \mathcal{A}} |\lambda_\alpha|$. The fibres of the zippered rectangles map

$$\text{zip} : \mathcal{F} \rightarrow \mathcal{C}, \quad \mathcal{C} = \mathcal{C}(\mathfrak{R})$$

are almost everywhere finite with constant cardinality depending on \mathcal{C} . Henceforth a superscript (1) means the version of the object where the corresponding translation surface has area 1: for example $\mathcal{F}^{(1)}, \mathcal{C}^{(1)}, X^{(1)}$. We also write $\mathcal{F}_\pi = \mathcal{F} \cap X_\pi$ and $\mathcal{F}_\gamma = \mathcal{F} \cap X_\gamma$.

2.10 The Teichmüller flow on suspension data

Recall that \mathcal{C} is a connected component of \mathcal{M}_κ and \mathfrak{R} the associated Rauzy class. The *Teichmüller flow* is a one parameter flow on $X_\mathfrak{R}$ that commutes with \hat{Q} and is given by $T_t(\pi, \lambda, \tau) = (\pi, e^t \lambda, e^{-t} \tau)$. Note that this preserves each $X_\pi^{(1)}$ and $X^{(1)}$. The flow T_t lifts the Teichmüller flow on \mathcal{C} , that is,

$$\mathcal{T}_t \circ \text{zip} = \text{zip} \circ T_t.$$

Evidently, T_t preserves Lebesgue measure on X . The flow T_t also preserves Lebesgue measure on $X^{(1)}$, the pushforward of which under zip is a multiple of $\nu_{\mathcal{C}(1)}$.

2.11 Time acceleration and renormalization.

The approach of Avila, Gouëzel and Yoccoz [3] to the Teichmüller flow is to consider the first return time to an appropriately chosen cross section. This cross section involves the choice of $\pi \in \mathfrak{R}$ and a path γ_0 that begins and ends at π . We give details on the choice of γ_0 in Section 2.13 and 4.1. For now, assume we have chosen π and γ_0 .

We consider the regions

$$\hat{\Xi} := \{ (\pi, \lambda, \tau) \in \mathcal{F}_{\gamma_0}^{(1)} : \|\lambda\| = 1 \} \cap (\{\pi\} \times Y_{\gamma_0} \times \mathcal{K}_{\gamma_0})$$

and the closely related

$$\Xi := \{ (\pi, \lambda) \in Y_{\gamma_0} : \|\lambda\| = 1 \}.$$

Let \hat{m} (resp. m) denote the normalized natural Lebesgue measure on $\hat{\Xi}$ (resp. Ξ). It is known that almost all orbits of the Teichmüller flow pass through $\hat{Q}^{\mathbf{Z}}(\Xi)$, this is stated in [3, 4.1.3] as a consequence of the ergodicity of the Veech flow¹¹. For each $x \in \hat{\Xi}$ we denote by $r(x)$ the first return time of x to $\hat{Q}^{\mathbf{Z}}(\Xi)$ under the Teichmüller flow. That is, $r(x)$ is the smallest positive value such that

$$T_{r(x)}(x) \in \hat{Q}^{-n}(\hat{\Xi})$$

for some positive¹² integer n . This means there is some value $\hat{Z}(x) \in \hat{\Xi}$ such that

$$T_{r(x)}\hat{Q}^n(x) = \hat{Q}^n T_{r(x)}(x) = \hat{Z}(x). \quad (2.11)$$

Suppose that $x = (\pi, \lambda, \tau) \in X_\gamma$ with $\hat{Q}^n(x) = (\pi, (\Theta_\gamma^*)^{-1}\lambda, (\Theta_\gamma^*)^{-1}\tau) \in \hat{\Xi}$. Then

$$r(x) = -\log \|(\Theta_\gamma^*)^{-1}\lambda\|.$$

Note here that $r(\pi, \lambda, \tau)$ depends only on the coordinates (π, λ) and we can view r also as a function on Ξ .

We will write $\gamma_1.\gamma_2$ or just $\gamma_1\gamma_2$ for the concatenation of two oriented paths γ_1 and γ_2 in \mathfrak{R} with compatible endpoints. In $\gamma_1.\gamma_2$, γ_1 is the first path traversed. Consider γ with the property that the γ_0 subpaths of $\gamma.\gamma_0$ are precisely the beginning and the end segment. We say that such a γ is γ_0 -adapted. For such a γ , if $x \in X_{\gamma.\gamma_0} \cap \hat{\Xi}$ then

$$\hat{Z}(x) = \left(\pi, \frac{(\Theta_\gamma^*)^{-1}\lambda}{\|(\Theta_\gamma^*)^{-1}\lambda\|}, \|(\Theta_\gamma^*)^{-1}\lambda\|(\Theta_\gamma^*)^{-1}\tau \right).$$

The domain of \hat{Z} is therefore $\cup_{\gamma_0\text{-adapted } \gamma} \hat{\Xi}_{\gamma\gamma_0}$ where

$$\hat{\Xi}_{\gamma\gamma_0} := \hat{\Xi} \cap (Y_{\gamma\gamma_0} \times \mathcal{K}_{\gamma_0}).$$

We extend this definition to $\hat{\Xi}_{\gamma_1\dots\gamma_N\gamma_0} := \hat{\Xi} \cap (Y_{\gamma_1\dots\gamma_N\gamma_0} \times \mathcal{K}_{\gamma_0})$ where $\gamma_1, \dots, \gamma_N$ are a sequence of γ_0 -adapted paths with both endpoints equal to π .

Notice that the mapping \hat{Z} has the following properties.

1. \hat{Z} is a skew extension of the mapping $Z : \Xi \rightarrow \Xi$ defined Lebesgue almost everywhere by

$$Z(\pi, \lambda) = \left(\pi, \frac{(\Theta_\gamma^*)^{-1}\lambda}{\|(\Theta_\gamma^*)^{-1}\lambda\|} \right), \quad (\pi, \lambda) \in Y_{\gamma.\gamma_0}.$$

The connected components of the domain of Z are the sets

$$\Xi_{\gamma\gamma_0} := \Xi \cap Y_{\gamma\gamma_0}.$$

2. The maps \hat{Z} and Z preserve $\|\lambda\| = 1$. This is usually referred to as *renormalization*.

¹¹The Veech flow is not discussed in the current paper.

¹²Notice that from (2.7) that Θ does not decrease norms, so if $(\pi', \lambda', \tau') = \hat{Q}(\pi, \lambda, \tau)$ then $\|\lambda'\| \leq \|\lambda\|$.

3. The maps \hat{Z} and Z involve many iterations of Rauzy induction and this is usually referred to as *time acceleration*. This is first due to Zorich [34], see also [36, Section 5.3] for further discussion.
4. \hat{Z} (resp. Z) preserves the Lebesgue measure \hat{m} (resp. m).

Following [3, Section 4.2.1], in order to enforce hyperbolicity of the map \hat{Z} (cf. Proposition 3.1 and Lemma 3.2) one puts adapted metrics on Ξ and $\hat{\Xi}$. On Ξ we put the Hilbert metric d_Ξ coming from the inclusion $\Xi \rightarrow Y_\pi$ and on $\hat{\Xi}$ we consider the product metric

$$d_{\hat{\Xi}}((\pi, \lambda, \tau), (\pi, \lambda', \tau')) := d_\Xi((\pi, \lambda), (\pi, \lambda')) + d_{\mathcal{K}_\pi}(\tau, \tau')$$

where $d_{\mathcal{K}_\pi}$ is the Euclidean distance in \mathcal{K}_π . These metrics induce Finsler metric structures on Ξ and $\hat{\Xi}$ that make them into complete Finsler manifolds.

2.12 Flow on sections of associated bundles in the suspension model

We may now map

$$\hat{\Xi}_r := \{(x, s) : x \in \hat{\Xi}, s \in [0, r(x))\}$$

homeomorphically to a part of $X_\pi^{(1)}$ by the map

$$P : (x, s) \mapsto T_s(x). \quad (2.12)$$

The image $X_\pi'^{(1)}$ of P is up to a Lebesgue-null set, a fundamental domain for the action of \hat{Q} on $X^{(1)}$. Given a section of $H_1(\mathcal{C}^{(1)}; \rho)$, its pull back to $X^{(1)}$ is therefore determined (up to zero measure set) by its values on $X_\pi'^{(1)} \subset X_\pi^{(1)}$.

The pushforward of Lebesgue measure under the mapping in (2.12) is Lebesgue measure. We write $\hat{m}_r = \hat{m} \otimes \text{Leb}$ for the Lebesgue measure on $\hat{\Xi}_r$.

As explained in Section 2.3, \mathcal{T}_t acts by \mathcal{T}_t^* on sections of $H_1(\mathcal{C}^{(1)}; \rho)$. If (after pullback) we view a section $\tilde{\sigma}$ as a V -valued function satisfying (2.10) and then view $\tilde{\sigma}$ as a V -valued function $\hat{\sigma}$ on $\hat{\Xi}_r$ by the mapping in (2.12) then the action of T_t^* on $\hat{\sigma}$ will be denoted by \hat{T}_t^* and defined as follows. Let γ be γ_0 -adapted with $l(\gamma) = n$. If $x \in X_{\gamma, \gamma_0}' \cap \hat{\Xi}$ and $t + s \in [r(x), r(x) + r(\hat{Z}(x)))$ then

$$\begin{aligned} [\hat{T}_t^* \hat{\sigma}](x, s) &= \hat{\sigma}(x, t + s) \\ &= \tilde{\sigma}(T_{t+s}x) \stackrel{(2.10)}{=} \rho(\Theta_\gamma^*) \tilde{\sigma}(\hat{Q}^n T_{t+s}x) \\ &\stackrel{(2.11)}{=} \rho(\Theta_\gamma^*) \tilde{\sigma}(T_{t+s-r(x)} \hat{Z}(x)) = \rho(\Theta_\gamma^*) \hat{\sigma}(\hat{Z}(x), t + s - r(x)). \end{aligned}$$

Let

$$r^{(N)}(x) := r(x) + r(\hat{Z}(x)) + \dots + r(\hat{Z}^{N-1}(x)).$$

For $\gamma_1, \gamma_2, \dots, \gamma_N$ each γ_0 -adapted, $t + s \in [r^{(N)}(x), r^{(N+1)}(x))$ and $x \in X_{\gamma_1, \gamma_2, \dots, \gamma_N, \gamma_0}' \cap \hat{\Xi}$ we have then

$$[\hat{T}_t^* \hat{\sigma}](x, s) = \rho(\Theta_{\gamma_1}^*) \rho(\Theta_{\gamma_2}^*) \dots \rho(\Theta_{\gamma_N}^*) \cdot \hat{\sigma}(\hat{Z}^N(x), t + s - r^{(N)}(x)). \quad (2.13)$$

This is the master equation for the Teichmüller flow on sections of $H_1(\mathcal{C}^{(1)}; \rho)$ in our suspension model. Notice that the argument of $\hat{\sigma}$ in the right hand side of (2.13) defines a mapping we call

$$\hat{T}_t : \hat{\Xi}_r \rightarrow \hat{\Xi}_r, \quad \hat{T}_t(x, s) := (\hat{Z}^N(x), t + s - r^{(N)}(x))$$

for $x \in X'_{\gamma_1, \gamma_2, \dots, \gamma_N, \gamma_0} \cap \hat{\Xi}$ and $t + s \in [r^{(N)}(x), r^{(N+1)}(x))$. Then \hat{T}_t is the suspension flow over \hat{Z} with roof function r . The flow \hat{T}_t lifts the Teichmüller flow under the mapping in (2.12) and as a consequence, Lebesgue measure \hat{m}_r on $\hat{\Xi}_r$ is invariant under \hat{T}_t .

Since the roof function r depends only on a coordinate in Ξ we may also define

$$\Xi_r = \{(y, s) : y \in \Xi, s \in [0, r(y))\}.$$

We write m_r for the Lebesgue measure on Ξ_r . We also define for $r \in Z^{-(N-1)}(\Xi)$

$$r^{(N)}(y) := r(y) + r(Z(y)) + \dots + r(Z^{N-1}(y)).$$

We may define a similar operator to \hat{T}_t^* that we will call T_t^* that will act on V -valued functions on Ξ_r . For $\sigma : \Xi_r \rightarrow V$, $\gamma_1, \gamma_2, \dots, \gamma_N$ each γ_0 -adapted, $t + s \in [r^{(N)}(y), r^{(N+1)}(y))$ and $y \in \Xi_{\gamma_1, \gamma_2, \dots, \gamma_N, \gamma_0}$ we define

$$[T_t^* \sigma](y, s) = \rho(\Theta_{\gamma_1}^*) \rho(\Theta_{\gamma_2}^*) \dots \rho(\Theta_{\gamma_N}^*) \cdot \sigma(Z^N(y), t + s - r^{(N)}(y)). \quad (2.14)$$

We give Ξ_r and $\hat{\Xi}_r$ Finsler metrics that are the product of the Finsler metric on Ξ (resp. $\hat{\Xi}$) with the usual metric in the s direction.

2.13 Preliminary choice of γ_0 .

Recall γ_0 is a path in \mathfrak{A} beginning and ending in π . We now explain the choice of γ_0 that is made in [3]. Avila, Gouëzel and Yoccoz require that

(Strongly Positive) γ_0 is a *strongly positive* path, meaning that all the entries of Θ_{γ}^* are positive and moreover $(\Theta_{\gamma_0}^*)^{-1}$ maps $\bar{\mathcal{K}}_{\pi} - \{0\}$ into \mathcal{K}_{π} .

(Neat) γ_0 is *neat*, meaning that $\gamma_0 = \gamma' \gamma_e = \gamma_s \gamma'$ implies γ' is trivial or $\gamma' = \gamma_0$. This means in any path, occurrences of γ_0 are (edge) disjoint. Therefore γ_0 -adapted γ are precisely those of the form

$$\gamma = \gamma_0 \gamma'.$$

where γ' does not contain γ_0 as a subpath.

According to [3, Section 4.13], such a choice of γ_0 is possible. However, in the present paper, we must choose γ_0 more carefully, while still making sure γ_0 is strongly positive and neat. This is done in Section 4.1. For now, assume that γ_0 is strongly positive and neat.

3 Decay of correlations

In this section we state in more precise terms and then prove Theorem 1.6 on uniform exponential decay of correlations.

3.1 Dynamical setup

The following definitions and results are from [3]. Recall the maps \hat{Z} and Z introduced in Section 2.11. Throughout we use the Finsler metric on the tangent bundle to Ξ defined in Section 2.11. We write D for the total derivative of a function. We write $C^0(\Xi)$ for the uniform norm. For a V -valued function F , $\|DF\|$ refers to operator norm w.r.t. the Finsler metric on the fibres and the Hilbert space metric on V . When we write \bigcup_γ^* or \sum_γ^* it means that we restrict the indexing to γ_0 -adapted γ . We assume here that γ_0 is strongly positive and neat as in Section 2.13, since these are required for the results of Avila, Gouëzel and Yoccoz [3].

Proposition 3.1 ([3, Proof of Proposition 4.3]). *The map Z is a **uniformly expanding Markov map** with respect to Lebesgue measure m and the Finsler metric structure defined in Section 2.11. That is to say*

1. *The union*

$$\bigcup_\gamma^* \Xi_{\gamma\gamma_0}$$

is a countable union of open sets that are m -conull in Ξ .

2. *If γ is γ_0 -adapted, Z maps $\Xi_{\gamma\gamma_0}$ diffeomorphically to Ξ and there are constants $\Lambda > 1$ and $c_1(\gamma) > 0$ such that for all $x \in \Xi_{\gamma\gamma_0}$ and v in the tangent fibre to x*

$$\Lambda\|v\| \leq \|[DZ]_x.v\| \leq c_1(\gamma)\|v\|.$$

3. *Letting J denote the inverse of the Jacobian of Z with respect to m . The function $\log J$ is C^1 on each $\Xi_{\gamma\gamma_0}$ and there is some $C > 0$ such that for any inverse branch α of Z ,*

$$\sup_{y \in \Xi} \|D(\log J \circ \alpha)(y)\| \leq C.$$

Lemma 3.2 ([3, Lemma 4.3]). *The pair (\hat{Z}, \hat{m}) is a **hyperbolic skew product** over (Z, m) . This means, with all norms and distances coming from the Finsler metric on $\hat{\Xi}$ defined in Section 2.11,*

1. *The projection $\text{pr} : \hat{\Xi} \rightarrow \Xi$ defined by*

$$\text{pr}(\pi, \lambda, \tau) = (\pi, \lambda)$$

satisfies $Z \circ \text{pr} = \text{pr} \circ \hat{Z}$ whenever both sides of the equality are defined.

2. *The measure \hat{m} gives full mass to the domain of definition of \hat{Z} .*
3. *There is a family of probability measures $\{\hat{m}_y\}_{y \in \Xi}$ on $\hat{\Xi}$ which is a disintegration of \hat{m} over m in the following sense: $y \mapsto \hat{m}_y$ is measurable, \hat{m}_y is supported on $\text{pr}^{-1}(y)$ and for any measurable $U \subset \hat{\Xi}$, $\hat{m}(U) = \int_{y \in \Xi} \hat{m}_y(U) dm(y)$. Moreover, there is a constant $C > 0$ such that for any open $V \subset Z^{-1}(\Xi)$, for any $u \in C^1(\text{pr}^{-1}(V))$ the function $\bar{u}(x) = \int u(x) d\hat{m}_y(x)$ is in $C^1(V)$ with*

$$\sup_{y \in V} \|D\bar{u}(x)\| \leq C \sup_{x \in \text{pr}^{-1}(V)} \|Du(y)\|.$$

4. There is a constant $K > 1$ such that for all $x_1, x_2 \in \hat{\Xi}$ with $\text{pr}(x) = \text{pr}(y)$ we have

$$d_{\hat{\Xi}}(\hat{Z}(x_1), \hat{Z}(x_2)) \leq K^{-1} d_{\hat{\Xi}}(x_1, x_2).$$

Lemma 3.3 ([3, Lemma 4.5]). *The roof function r is **good**. This means*

1. There is $\epsilon_1 > 0$ such that $r \geq \epsilon_1$.
2. There is $C > 0$ such that for any inverse branch α of Z one has

$$\sup_{y \in \Xi} \|D(r \circ \alpha)(y)\| \leq C.$$

3. There is no C^1 function ϕ on $\bigcup_{\gamma}^* \Xi_{\gamma\gamma_0}$ such that

$$r - \phi \circ T + \phi$$

is constant on each $\Xi_{\gamma\gamma_0}$.

Theorem 3.4 ([3, Theorem 4.6]). *The roof function r has **exponential tails**. This means there is $\sigma_0 > 0$ such that*

$$\int_{\Xi} \exp(\sigma_0 r) dm < \infty.$$

3.2 The main technical results

The following will be the precise version of Theorem 1.6. Recall the definition of $\mathcal{C}^{(1)}(q)$ from Section 2.3.

Theorem 3.5. *Let $\mathfrak{R} = \mathfrak{R}(\mathcal{C})$ and $\pi \in \mathfrak{R}$. Assume that G_{π} is Zariski-dense. There is some $\eta, C, \delta > 0$ and integer $q_0 > 0$ such that for q coprime to q_0 , for all compactly supported $u, v \in C^1(\mathcal{C}^{(1)}(q)) \cap L^2_0(\mathcal{C}^{(1)}(q))$ and all $t \geq 0$*

$$\left| \int u \cdot v \circ \mathcal{T}_t d\nu_{\mathcal{C}^{(1)}(q)} - \left(\int u d\nu_{\mathcal{C}^{(1)}(q)} \right) \left(\int v d\nu_{\mathcal{C}^{(1)}(q)} \right) \right| \leq C \|u\|_{C^1} \|v\|_{C^1} q^{\eta} e^{-\delta t}.$$

If G_{π} is not finite index in $\text{Sp}(\mathbf{Z}^{2g}, \omega_{\pi})$ we must also assume q is squarefree.

For any Finsler manifold X and Hilbert space V we may define the Banach space of C^1 V -valued functions on X as in Section 2.3. Recall from Sections 2.11 and 2.12 that there are Finsler metric structures on $\Xi, \hat{\Xi}, \Xi_r, \hat{\Xi}_r$. If (ρ, V) is a unitary representation we write e.g. $C^1(\Xi; \rho)$ for the C^1 V -valued functions on Ξ , with respect to the Finsler metric. We make a reduction of Theorem 3.5 to the following that is analogous to [3, Theorem 2.7].

Theorem 3.6. *Make the same assumptions on \mathcal{C} and q as in Theorem 3.5. Then there is some $\eta, C, \delta > 0$ and integer $q_0 > 0$ such that for q coprime to q_0 , we have for all $U, V \in C^1(\hat{\Xi}_r; \rho_q)$ and all $t \geq 0$*

$$\left| \int \langle U, T_t^* V \rangle d\hat{m}_r - \left\langle \int U d\hat{m}_r, \int V d\hat{m}_r \right\rangle \right| \leq C \|U\|_{C^1} \|V\|_{C^1} q^{\eta} e^{-\delta t}.$$

After applying the correspondences of Lemma 2.1 and Section 2.12, the lifting and smoothing argument that converts Theorem 3.6 to Theorem 3.5 appears in [3, pp. 166-169]. It is technical and relies on the hard Theorem 3.4, as well as comparisons between Finsler metric structures. However, it applies as well to vector valued functions as scalar valued functions and the constants involved have no dependence on the vector space so it goes through in the current setting.

3.3 Entrance of the transfer operator

We now recall the definition of the spaces \mathcal{B}_0 and \mathcal{B}_1 from [3].

Definition 3.7. A function $U : \Xi_r \rightarrow V$ is in $\mathcal{B}_0(\Xi_r; \rho)$ if it is bounded, continuously differentiable on each set

$$(\Xi_r)_{\gamma\gamma_0} := \{(y, t) : y \in \Xi_{\gamma\gamma_0}, t \in (0, r(y))\} \quad \gamma \text{ is } \gamma_0\text{-adapted}$$

and also $\sup_{(y,t) \in \cup^*(\Xi_r)_{\gamma\gamma_0}} \|DU(y, t)\| < \infty$. Define the norm

$$\|U\|_{\mathcal{B}_0(\Xi_r; \rho)} := \sup_{(y,t) \in \cup^*(\Xi_r)_{\gamma\gamma_0}} \|U(y, t)\| + \sup_{(y,t) \in \cup^*(\Xi_r)_{\gamma\gamma_0}} \|DU(y, t)\|.$$

Definition 3.8. A function $U : \Xi_r \rightarrow V$ is in $\mathcal{B}_1(\Xi_r; \rho)$ if it is bounded and there exists a constant $C > 0$ such that for all fixed $y \in \cup^*\Xi_{\gamma\gamma_0}$, the function $t \mapsto U(y, t)$ is of bounded variation¹³ on the interval $(0, r(y))$ and its variation $\text{Var}_{(0, r(y))}(t \mapsto U(y, t))$ is bounded by $Cr(y)$. Let

$$\|U\|_{\mathcal{B}_1} = \sup_{(y,t) \in \cup^*(\Xi_r)_{\gamma\gamma_0}} \|U(y, t)\| + \sup_{y \in \cup^*(\Xi)_{\gamma\gamma_0}} \frac{\text{Var}_{(0, r(y))}(t \mapsto U(y, t))}{r(y)}.$$

As in [3] we reduce to decay of correlations for the ρ -skew extension of Ξ_r rather than $\hat{\Xi}_r$.

Theorem 3.9 (Decay of correlations). *Make the same assumptions on \mathcal{C} and q as in Theorem 3.5. There exist constants $C > 0$ and $\delta > 0$ such that for all $U \in \mathcal{B}_0(\Xi_r; \rho_q)$ and $V \in \mathcal{B}_1(\Xi, \rho_q)$, for all $t \geq 0$,*

$$\left| \int \langle U, T_t^* V \rangle dm_r - \left\langle \int U dm_r, \int V dm_r \right\rangle \right| \leq C \|U\|_{\mathcal{B}_0} \|V\|_{\mathcal{B}_1} e^{-\delta t}.$$

This is proved for scalar valued functions in [3, Theorem 7.3]. The key point of Theorem 3.9 is the uniformity in q . The passage from Theorem 3.9 to Theorem 3.6 is handled as in [3, Section 8]. In fact, the arguments of [3, Section 8] are followed closely and extended to the skew setting by Oh and Winter in [26, Proof of Theorem 1.5]. So we have presently explained the reduction of Theorem 3.5 to Theorem 3.9 whose proof we now take up.

From now on, all integrals are taken with respect to the relevant Lebesgue measure. In proving Theorem 3.9 we may assume that $\int_{\Xi_r} V = 0$. Following [3] let

$$A_t = \{(y, a) \in \Xi_r : a + t \geq r(y)\}$$

and $B_t = \Xi_r \setminus A_t$. We bound

$$\int_{B_t} \langle U, T_t^* V \rangle \leq \|U\|_{\mathcal{B}_0} \|V\|_{\mathcal{B}_1} \int_{y \in \Xi} \max(r(y), 0) \leq \|U\|_{\mathcal{B}_0} \|V\|_{\mathcal{B}_1} \int_{y: r(y) \geq t} r(y).$$

By Cauchy-Schwarz inequality and that r has exponential tails (Theorem 3.4) the above contributes $\leq C' \|U\|_{\mathcal{B}_0} \|V\|_{\mathcal{B}_1} \exp(-\delta' t)$ for some $\delta' > 0$ and $C' > 0$ that do not depend on U, V or ρ . Therefore the proof of Theorem 3.9 reduces to estimating the quantity

¹³We make the obvious extension of bounded variation to V -valued functions using the norm induced by the inner product on the Hilbert space V .

$$I(t) := \int_{A_t} \langle U, T_t^* V \rangle$$

on the order of

$$I(t) \leq C \|U\|_{\mathcal{B}_0} \|V\|_{\mathcal{B}_1} \exp(-\delta t) \quad (3.1)$$

for some absolute constants $C, \delta > 0$.

We now begin the proof of (3.1). We will estimate the Laplace transform

$$\hat{I}(s) := \int_0^\infty \exp(-st) I(t) dt. \quad (3.2)$$

This is convergent for $\Re(s) > 0$ since I is bounded using the finiteness of m_r . The estimation of $\hat{I}_s(t)$ is closely related to certain skew transfer operators as follows. Using notation of [3], if $F : \Xi_r \rightarrow V$ and $s \in \mathbf{C}$, let

$$\hat{F}_s(y) := \int_0^{r(y)} F(y, \tau) \exp(-s\tau) d\tau.$$

Then following the proof of [3, Lemma 7.17] and adapting to our ρ -skew setting we have

$$\begin{aligned} \hat{I}(s) &= \int_{y \in \Xi} \int_{\tau=0}^{r(y)} \int_{t+\tau \geq r(y)} e^{-st} \langle U(y, \tau), [T_t^* V](y, \tau) \rangle dt d\tau dy \\ &= \sum_{k=1}^\infty \int_{y \in \Xi} \int_{\tau=0}^{r(y)} \int_{\tau'=0}^{r(Z^k y)} e^{-s(r^{(k)}(y) + \tau' - \tau)} \langle U(y, \tau), [T_t^* V](y, \tau) \rangle d\tau' d\tau dy. \end{aligned} \quad (3.3)$$

The manipulation above follows from writing for each y , $t + \tau = r^{(k)}(y) + \tau'$ with $\tau' \in [0, r(Z^k x))$. For each y and t there is a unique k and τ' for which this is possible. Supposing more specifically that $y \in \Xi_{\gamma_1, \dots, \gamma_k \gamma_0}$ with each γ_i γ_0 -adapted, we get from (2.14) that

$$[T_t^* V](y, \tau) = \rho(\Theta_{\gamma_1}^*) \rho(\Theta_{\gamma_2}^*) \dots \rho(\Theta_{\gamma_k}^*) \cdot V(Z^k(y), \tau'). \quad (3.4)$$

Inserting this into (3.3) gives that (throwing out a measure zero set)

$$\begin{aligned} \hat{I}(s) &= \sum_{k=1}^\infty \sum_{\gamma_1, \dots, \gamma_k}^* \int_{y \in \Xi_{\gamma_1, \dots, \gamma_k \gamma_0}} \int_{\tau=0}^{r(y)} \int_{\tau'=0}^{r(Z^k y)} e^{-s(r^{(k)}(y) + \tau' - \tau)} \langle U(y, \tau), (3.4) \rangle d\tau' d\tau dy \\ &= \sum_{k=1}^\infty \sum_{\gamma_1, \dots, \gamma_k}^* \int_{y \in \Xi_{\gamma_1, \dots, \gamma_k \gamma_0}} e^{-sr^{(k)}(y)} \langle \hat{U}_{-s}(y), \rho(\Theta_{\gamma_1}^*) \rho(\Theta_{\gamma_2}^*) \dots \rho(\Theta_{\gamma_k}^*) \hat{V}_s(Z^k(y)) \rangle dy. \end{aligned} \quad (3.5)$$

Here, we write a \sum^* to indicate that the γ_i being summed over are all γ_0 -adapted. The expression (3.5) is best understood by the *skew transfer operator* that we now introduce. Recall that $y \in \Xi$ can be written $y = (\pi, \lambda)$. The inverse branches of Z are indexed by γ_0 -adapted γ and are given explicitly by

$$\alpha_\gamma : (\pi, \lambda) \mapsto \left(\pi, \frac{\Theta_\gamma^* \lambda}{\|\Theta_\gamma^* \lambda\|} \right), \quad \Xi \rightarrow \Xi_{\gamma_0}. \quad (3.6)$$

The skew transfer operator $\mathcal{L}_{s,\rho}$ is defined for arbitrary unitary (ρ, V) and $f : \Xi \rightarrow V$ by

$$\mathcal{L}_{s,\rho}[f](y) := \sum_{\gamma}^* e^{-sr \circ \alpha_{\gamma}(y)} J \circ \alpha_{\gamma}(y) \rho(\Theta_{\gamma}^*)^{-1} \cdot f \circ \alpha_{\gamma}(y).$$

Recall that J is the inverse of the Jacobian of Z w.r.t. Lebesgue measure. By results of [3] the summation involved in $\mathcal{L}_{s,\rho}$ is absolutely uniformly convergent on compact subsets of $\cup^* \Xi_{\gamma\gamma_0}$ (cf. Theorem 4.4 and the discussion afterwards). With the operator $\mathcal{L}_{s,\rho}$ in hand, by making a change of variables of the form $y \mapsto Z^k(y)$ one obtains from (3.5)

$$\hat{I}(s) = \sum_{k=1}^{\infty} \int_{y \in \Xi} \langle \mathcal{L}_{s,\rho}^k[\hat{U}_{-s}](y), \hat{V}_s(y) \rangle dy. \quad (3.7)$$

It is clear from inspection of the above that spectral bounds for the operator $\mathcal{L}_{s,\rho}$ will be helpful in estimating \hat{I} .

3.4 Spectral bounds for transfer operators.

a. $|\Im(s)| \gg 1$. Here we give spectral bounds for transfer operators $\mathcal{L}_{s,\rho}$, where ρ is an arbitrary unitary representation, that come from the method of Dolgopyat [10]. In the case of scalar valued functions on Ξ these bounds were obtained by Avila, Gouëzel and Yoccoz in [3] by adapting Dolgopyat's argument to the Teichmüller setting. As has been shown by Oh and Winter [26] (see also [20] for this argument in another setting) Dolgopyat's argument functions perfectly well for skew transfer operators, provided the twisting unitary cocycle is constant on cylinders of length 1. In the current setting, this is true since the values of the cocycle

$$\rho(\Theta_{\gamma})$$

only depend on the cylinder $\Xi_{\gamma\gamma_0}$. To keep the length of the paper down, and to avoid repeating parts of either [26] or [20], we simply state the results we need and refer the reader to (*loc. cit.*) for the necessary arguments.

To state the next result we introduce the warped norm on $C^1(\Xi; \rho)$ by

$$\|u\|_{1,t} = \sup_{y \in \Xi} \|u(y)\| + \frac{1}{\max(1, |t|)} \sup_{y \in \Xi} \|Du(y)\|.$$

The main result we wish to state in this section is

Proposition 3.10. *There is $\sigma'_0 \leq \sigma_0$, $T_0 > 0$, $C > 0$ and $\beta < 1$ such that for all $s = \sigma + it$ with $|\sigma| \leq \sigma'_0$ and $|t| \geq T_0$, for any unitary (ρ, V) , $u \in C^1(\Xi; \rho)$ and for all $k \in \mathbf{N}$*

$$\|\mathcal{L}_{s,\rho}^k u\|_{L^2(\Xi)} \leq C \beta^k \|u\|_{1,t}.$$

The version of Proposition 3.10 with no twist by ρ can be found in [3, Proposition 7.7]. See Oh and Winter [26, Theorem 3.1] for the extension to skew operators and [20, Lemma 23] for another version of this type of estimate.

b. $|\Im(s)| \ll 1$. Here we give spectral bounds for \mathcal{L}_{s,ρ_q} that are good when $|\Im(s)|$ is below a fixed constant.

Proposition 3.11. *Assume G_π is Zariski-dense in $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_\pi)$. Let $s = \sigma + it$. There is a positive integer q_0 such that for all $t_0 > 0$ there are constants $c, C, \eta, \epsilon > 0$ and $0 < \sigma_1 < \sigma_0$ such that when $|\sigma| < \sigma_1$ and $|t| < t_0$ then for all $u \in C^1(\Xi; \rho)$, all $k \in \mathbf{N}$ and q coprime to q_0 we have*

$$\|\mathcal{L}_{s, \rho_q}^k u\|_{C^1} \leq C(1 - \epsilon)^k q^\eta \|u\|_{C^1}.$$

If G_π is Zariski-dense but not finite index in $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_\pi)$ we must also assume that q is squarefree.

Propositions 3.10 and 3.11 together with the expression (3.7) imply Theorem 3.9 by the arguments of [26, Lemma 5.3, Lemma 5.4, Proposition 5.5]. The only outstanding proof required for Theorem 3.5 is Proposition 3.11, this is given in Section 4.

4 Expansion and the twisted transfer operator

This section contains a proof of Proposition 3.11.

4.1 Refining the choice of γ_0

Recall the fixed member π of a Rauzy class \mathfrak{R} , and the definition of the Rauzy-Veech group G_π from Section 2.7. We assume that one of the following scenarios occurs.

- A.** The group G_π is finite index in $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_\pi)$.
- B.** The group G_π is Zariski-dense in $\mathrm{Sp}(\mathbf{C}^{2g}, \omega_\pi)$.

We prefer to distinguish scenario **A** if it is available since it implies **B** by the Borel density theorem [7] but has stronger consequences. Note that in either case there is a finitely generated subgroup G'_π with the same property as G_π : in case **A** set $G'_\pi = G_\pi$ and in case **B** let G'_π be a Zariski-dense finitely generated subgroup of G_π . Let S denote the finite set of generators of G'_π . Choose a finite set Υ_0 of γ that are paths in \mathfrak{R} beginning and ending in π and such that

$$\{\Theta_\gamma^* : \gamma \in \Upsilon_0\}$$

together with their inverses generate S . Now let

$$\Upsilon = \Upsilon_0 \cup \{\gamma \cdot \gamma : \gamma \in \Upsilon_0\}.$$

While this might seem a bit mysterious, we note for later that this definition guarantees

Lemma 4.1. *The elements*

$$\Theta_\gamma^* \cdot (\Theta_{\gamma'}^*)^{-1} \in \mathrm{Sp}(\mathbf{Z}^{2g}, \omega_\pi) \quad \gamma, \gamma' \in \Upsilon$$

generate G'_π .

Proof. For a given $\tilde{\gamma} \in \Upsilon_0$, we have $\Theta_{\tilde{\gamma} \cdot \tilde{\gamma}}^* (\Theta_{\tilde{\gamma}}^*)^{-1} = \Theta_{\tilde{\gamma}}^*$ and $\Theta_{\tilde{\gamma}}^* (\Theta_{\tilde{\gamma} \cdot \tilde{\gamma}}^*)^{-1} = (\Theta_{\tilde{\gamma}}^*)^{-1}$. On the other hand, the $\Theta_{\tilde{\gamma}}^*$ with $\tilde{\gamma} \in \Upsilon_0$ together with their inverses generate S and hence G'_π . \square

We will now choose γ_0 such that no $\gamma \in \Upsilon$ contains γ_0 as a substring and moreover γ_0 is strongly positive and neat (recall these properties from Section 2.13). This can be done simply by ensuring that γ_0 is strongly positive and neat and longer than all $\gamma \in \Upsilon$.

Before stating the next lemma we introduce some language. A path in \mathfrak{R} is *complete* if every $\alpha \in \mathcal{A}$ is the winner of some arrow in γ . It follows from a result of [22, Section 1.2.3] that there exists a complete path γ_* beginning and ending at π . A path in \mathfrak{R} is said to be *k-complete* if it is the concatenation of k complete paths. Write γ_*^k for the k -fold concatenation of γ_* with itself. Then for example, if γ_* is complete then γ_*^k is k -complete.

Lemma 4.2 ([3, Lemma 4.2]). *A k -complete path with $k \geq 3|\mathcal{A}| - 4$ is strongly positive.*

As noted in [3, pg. 162, footnote], a path is neat if it ends with a type ϵ arrow and begins with a string of opposite type arrows at least half the length of the path. Suppose that γ_* ends with a bottom arrow. Choose then k such that

$$l(\gamma_*) \cdot k \geq \max_{\gamma \in \Upsilon} l(\gamma), \quad k \geq 3|\mathcal{A}| - 4. \quad (4.1)$$

Next choose γ' beginning and ending at π with $l(\gamma_*) \cdot k + |\mathfrak{R}|$ top arrows at its beginning and $\leq |\mathfrak{R}|$ arrows afterwards (this is always possible since whatever the endpoint of the first top arrows, one can quickly return to π). Then

$$\gamma_0 := (\gamma' \gamma_*) \gamma_*^{k-1}$$

begins with

$$l(\gamma_*) \cdot k + |\mathfrak{R}| = \frac{1}{2} (l(\gamma_*) \cdot k + |\mathfrak{R}| + |\mathfrak{R}| + l(\gamma_*) \cdot k) \geq \frac{1}{2} l(\gamma_0)$$

top arrows so is therefore neat. Also, clearly $\gamma' \gamma_*$ is complete so γ_0 is k -complete. Therefore γ_0 is strongly positive by Lemma 4.2. Finally, by choice of k in (4.1) γ_0 is longer than any element of Υ . We have shown

Lemma 4.3. *It is possible to choose γ_0 so that no element $\gamma \in \Upsilon$ contains γ_0 as a substring and moreover γ_0 is strongly positive and neat.*

We fix such a γ_0 for the remainder of the paper (and retroactively for the previous sections). From the discussion in Section 2.13 this has the consequence that the elements of the set

$$\gamma_0 \cdot \Upsilon := \{ \gamma_0 \gamma : \gamma \in \Upsilon \}$$

are all γ_0 -adapted. We will use this later.

4.2 Decoupling I: Releasing the convolution

We now perform the decoupling argument begun in [20] with the central part of the argument coming from [20, Appendix]. One key difference here is the fact that the symbolic dynamics takes place on an infinite alphabet.

For $\gamma_1, \dots, \gamma_k$ γ_0 -adapted let $\alpha_{\gamma_1 \dots \gamma_k}$ denote the inverse branch of Z^k that maps Ξ to $\Xi_{\gamma_1 \dots \gamma_k \gamma_0}$. Then recalling the previously defined α_γ from (3.6) one has the composition law

$$\alpha_{\gamma_1 \dots \gamma_k} = \alpha_{\gamma_1} \circ \alpha_{\gamma_2} \circ \dots \circ \alpha_{\gamma_k}.$$

We write out to begin with, for $F \in C^1(\Xi; \rho)$,

$$\mathcal{L}_{s,\rho}^N[F](y) = \sum_{\gamma_1, \dots, \gamma_N} e^{[-sr^{(N)} + (\log J)^{(N)}](\alpha_{\gamma_1 \dots \gamma_N} y)} \rho(\Theta_{\gamma_N}^*)^{-1} \dots \rho(\Theta_{\gamma_1}^*)^{-1} F(\alpha_{\gamma_1 \dots \gamma_N} y),$$

where for $y' \in \Xi_{\gamma_1 \dots \gamma_N \gamma_0}$

$$(\log J)^{(N)}(y') := \log J(y') + \log J(Z(y')) + \dots + \log J(Z^{N-1}(y')).$$

It will be convenient to introduce the functions $r_s := -sr + \log J$ and $r_s^{(n)} := -sr^{(n)} + (\log J)^{(n)}$.

We understand for the rest of this section that all γ_i are γ_0 -adapted in all sums and so forth. We will frequently compare $\mathcal{L}_{s,\rho}$ to the operator on scalar functions

$$\mathcal{L}_s[f](y) := \sum_{\gamma} e^{-sr \circ \alpha_{\gamma}(y)} J \circ \alpha_{\gamma}(y) f(\alpha_{\gamma}(y))$$

that features in [3, formula (7.13)]. Recall that σ_0 is such that $\int \exp(\sigma_0 r) dm < \infty$ given by Theorem 3.4. The following is given in [3, pg. 188].

Theorem 4.4. *The operator \mathcal{L}_0 acts on $C^1(\Xi)$. There is some $0 < \sigma_1 < \sigma_0$ such that for s with $|\Re(s)| < \sigma_1$, \mathcal{L}_s is a bounded operator on $C^1(\Xi)$. Moreover we have the following properties after suitable choice of σ_1 :*

1. \mathcal{L}_0 has a simple eigenvalue at 1 and the rest of the spectrum of \mathcal{L}_0 is contained in a ball around 0 of radius < 1 .
2. For real σ with $|\sigma| < \sigma_1$ the leading eigenvalue λ_{σ} of \mathcal{L}_{σ} varies real analytically in σ . In particular for all $\eta > 0$ there is $\sigma_2(\eta) > 0$ such that for real σ with $|\sigma| \leq \sigma_2$ we have $\lambda_{\sigma} \leq e^{\eta}$.
3. The leading eigenfunctions h_{σ} (normalized so $\int h_{\sigma} = 1$) are positive and also vary real analytically as $C^1(\Xi)$ -valued functions on $(-\sigma_1, \sigma_1)$. The functions h_{σ} are uniformly bounded below when $|\sigma| \leq \sigma_1$.

As a corollary to Theorem 4.4 we may note that for real σ with $|\sigma| < \sigma_1$, the infinite sum

$$\sum_{\gamma} e^{r_{\sigma} \circ \alpha_{\gamma}(y)} = \mathcal{L}_{\sigma}[1](y)$$

converges to a C^1 function of $y \in \Xi$. Moreover it is possible to show by adapting the proof of [3, Lemma 7.8] that $\mathcal{L}_{s,\rho}$ acts on $C^1(\Xi, \rho)$ for $|\Re(s)| < \sigma_1$. We make the standing assumption in the following sections that $|\sigma| \leq \sigma_1$ from Theorem 4.4 so we can assume for example that $\|D(r_{\sigma} \circ \alpha_{\gamma})\|$ is bounded (independently of α_{γ}) on Ξ using Proposition 3.1 and Lemma 3.3. This implies that if $r_{\sigma}^{(M)}(x) := r_{\sigma}(x) + r_{\sigma}(Zx) + \dots + r_{\sigma}(Z^{M-1}x)$

$$\|Dr_{\sigma}^{(M)} \circ \alpha_{\gamma_1 \dots \gamma_M}\| \ll 1 \tag{4.2}$$

with constant independent of M and $\gamma_1, \dots, \gamma_M$ by summing the resulting geometric series.

Using that the roof function r is good (Lemma 3.3, Part 2 in particular) together with the bound on the derivative of $\log J$ (Proposition 3.1 Part 3) and the expanding property in

terms of constant $\Lambda > 1$ (Proposition 3.1 Part 2) it is possible to perform the same initial decoupling arguments as in [20, Section 5.1]. This yields after writing

$$N = M + \tilde{M}$$

and letting o be an arbitrary point in Ξ that

$$\begin{aligned} \mathcal{L}_{s,\rho}^N[F](y) &= \sum_{\gamma_1, \dots, \gamma_M} \sum_{\gamma_{M+1}, \dots, \gamma_N} \rho(\Theta_{\gamma_N}^*)^{-1} \dots \rho(\Theta_{\gamma_1}^*)^{-1} e^{r_s^{(N)}(\alpha_{\gamma_1 \dots \gamma_N} y)} F(\alpha_{\gamma_1 \dots \gamma_M} o) \\ &+ O(\|F\|_{C^1} \Lambda^{-M}) \\ &= \sum_{\gamma_1, \dots, \gamma_M} Op_{\gamma_1 \dots \gamma_M; y}(\rho) \cdot \rho(\Theta_{\gamma_M}^*)^{-1} \dots \rho(\Theta_{\gamma_1}^*)^{-1} F(\alpha_{\gamma_1 \dots \gamma_M} o) + O(\lambda_{\Re(s)}^N \|F\|_{C^1} \Lambda^{-M}) \end{aligned} \quad (4.3)$$

where

$$Op_{\gamma_1 \dots \gamma_M; y}(\rho) := \sum_{\gamma_{M+1}, \dots, \gamma_N} e^{r_s^{(N)}(\alpha_{\gamma_1 \dots \gamma_N} y)} \rho(\Theta_{\gamma_N}^*)^{-1} \dots \rho(\Theta_{\gamma_{M+1}}^*)^{-1}$$

is a member of the von Neumann algebra generated by the $\rho(\Theta_{\gamma_i}^*)^{-1}$ acting on V . Similarly following [20, Section 5.3] one may obtain that

$$\begin{aligned} D(\mathcal{L}_{s,\rho}^N[F])(y) &= \sum_{\gamma_1, \dots, \gamma_M} Op_{\gamma_1 \dots \gamma_M; y}^\partial(\rho) \cdot \rho(\Theta_{\gamma_M}^*)^{-1} \dots \rho(\Theta_{\gamma_1}^*)^{-1} F(\alpha_{\gamma_1 \dots \gamma_M} o) \\ &+ O(\lambda_{\Re(s)}^N (1 + \Im(s)) \|F\|_{C^1} \Lambda^{-M}) \end{aligned} \quad (4.4)$$

where

$$Op_{\gamma_1 \dots \gamma_M; y}^\partial(\rho) := \sum_{\gamma_{M+1}, \dots, \gamma_N} e^{r_s^{(N)}(\alpha_{\gamma_1 \dots \gamma_N} y)} D(r_s^{(N)} \circ \alpha_{\gamma_1 \dots \gamma_N})(y) \otimes \rho(\Theta_{\gamma_N}^*)^{-1} \dots \rho(\Theta_{\gamma_{M+1}}^*)^{-1}$$

is a member of $\text{Hom}(T_y \Xi, \mathbf{R}) \otimes \text{End}(V) \cong \text{Hom}(T_y \Xi, \text{End}(V))$ and the big O term is interpreted w.r.t. the operator norm between the Finsler metric norm on $T_y \Xi$ and $\text{End}(V)$ with its own operator norm. Write $\|\bullet\|_{T_y \Xi, \text{End}(V)}$ for this norm and $\|\bullet\|_{\text{End}(V)}$ for the operator norm on $\text{End}(V)$.

From now on we write $\sigma = \Re(s)$ and assume $|\sigma| \leq \sigma_1$ to be chosen smaller than the σ_1 of Theorem 4.4. Note that

$$\|\rho(\Theta_{\gamma_M}^*)^{-1} \dots \rho(\Theta_{\gamma_1}^*)^{-1} F(\alpha_{\gamma_1 \dots \gamma_M} o)\| \leq \|F\|_{C^1}$$

and a relatively straightforward operator norm bound for e.g. $Op_{\gamma_1 \dots \gamma_M; y}(\rho)$ is

$$\begin{aligned} \|Op_{\gamma_1 \dots \gamma_M; y}(\rho)\|_{\text{End}(V)} &\leq \sum_{\gamma_{M+1}, \dots, \gamma_N} e^{r_\sigma^{(N)}(\alpha_{\gamma_1 \dots \gamma_N} y)} \\ &= \sum_{\gamma_{M+1}, \dots, \gamma_N} e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_N} y)} e^{r_\sigma^{(\tilde{M})}(\alpha_{\gamma_{M+1} \dots \gamma_N} y)} \\ &= e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M} o)} (1 + O(\Lambda^{-M})) \mathcal{L}_\sigma^{\tilde{M}}[1](y) \end{aligned} \quad (4.5)$$

$$\ll e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M} o)} \lambda_\sigma^{\tilde{M}}. \quad (4.6)$$

Here we decoupled to get (4.5) by using $r_\sigma^{(N)}(\alpha_{\gamma_1 \dots \gamma_N} y) = r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M} o) + O(\Lambda^{-M})$. This bound can be obtained from the contracting properties of the α_{γ_i} and global bound for the derivative of $r_\sigma \circ \alpha_\gamma$. We will make this type of argument repeatedly. So the key point is to improve over (4.6) which is done in the following.

Proposition 4.5. *Make the same assumptions about G_π and q as in Proposition 3.11. Let $s = \sigma + it$. There is $q_0 \in \mathbf{N}_+$ and $D > 0$ such that for all $t_0 > 0$, there are $\sigma_1, c, C > 0$ such that for $|\sigma| < \sigma_1$, $|t| \leq t_0$, $(q, q_0) = 1$ and $\tilde{M} \approx c \log q$, we have*

$$\|Op_{\gamma_1 \dots \gamma_M; y}(\rho_q^{\text{new}})\|_{\text{End}(V)} \leq C e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M} o)} q^{-D/4} \lambda_\sigma^{\tilde{M}},$$

$$\|Op_{\gamma_1 \dots \gamma_M; y}^\partial(\rho_q^{\text{new}})\|_{T_y \Xi, \text{End}(V)} \leq C e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M} o)} q^{-D/4} \lambda_\sigma^{\tilde{M}}.$$

The bound for $Op_{\gamma_1 \dots \gamma_M; y}^\partial(\rho_q^{\text{new}})$ is similar to that for $Op_{\gamma_1 \dots \gamma_M; y}(\rho_q^{\text{new}})$ with essentially no added difficulties¹⁴, so we treat only $Op_{\gamma_1 \dots \gamma_M; y}(\rho_q^{\text{new}})$ to be efficient. The reader can consult [20, Section 5.3] for more details. The proof of Proposition 4.5 will take up the remaining Subsections 4.3, 4.4 of the present section.

The passage from Proposition 4.5 to Proposition 3.11 can be found in [9], see also [20, Section 5.5]. It requires possibly increasing q_0 and choosing σ_1 suitably small.

4.3 Bounding the operator norm of convolution operators

Recall $\Pi_q : \text{Sp}(\mathbf{Z}^{2g}, \omega_\pi) \rightarrow \Gamma_q$ is the reduction mod q map. We make the usual assumption that strong approximation holds at q and hence after suitable identification $\Gamma_q = \text{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_\pi)$.

To improve the readability of the following argument we will write for γ_0 -adapted γ

$$h_\gamma := \Pi_q(\Theta_\gamma^*)^{-1} \in \Gamma_q.$$

We are tasked with estimating the operator norm of the group algebra element

$$\mu_{\gamma_1 \dots \gamma_M; y} := \sum_{\gamma_{M+1}, \dots, \gamma_N} e^{r_s^{(N)}(\alpha_{\gamma_1 \dots \gamma_N} y)} h_{\gamma_N} h_{\gamma_{N-1}} \dots h_{\gamma_{M+1}} \in \mathbf{C}[\Gamma_q]$$

as it acts by convolution on $\ell^2(\Gamma_q)$. Indeed, this is precisely the operator $Op_{\gamma_1 \dots \gamma_M; y}(\rho_q^{\text{new}})$ when restricted to $\ell_{\text{new}}^2(\Gamma_q)$. We view elements of $\mathbf{C}[\Gamma_q]$ interchangeably as complex valued measures on Γ_q . We write $*$ for the convolution of measures, this corresponds to multiplication in $\mathbf{C}[\Gamma_q]$. Given $\mu \in \mathbf{C}[\Gamma_q]$ we write $|\mu|$ for the non negative real measure obtained by taking absolute values of coefficients.

Let $\sigma = \Re(s)$ and recall $N = M + \tilde{M}$. By a further decoupling argument (cf. [20, Lemma 38]) one obtains for arbitrary $o \in \Xi$

$$|\mu_{\gamma_1 \dots \gamma_M; y}| \leq C e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M} o)} \mu_1 \tag{4.7}$$

¹⁴ After applying $Op_{\gamma_1 \dots \gamma_M; y}^\partial(\rho_q^{\text{new}})$ to a test tangent vector v in $T_y \Xi$ one obtains an element of $\text{End}(V)$ with the task of bounding its operator norm, which can be done in exactly the same way as we will treat $Op_{\gamma_1 \dots \gamma_M; y}(\rho_q^{\text{new}})$. On the other hand, it is worth pointing out that the bound for $Op_{\gamma_1 \dots \gamma_M; y}^\partial(\rho_q^{\text{new}})$ relies crucially on the fact that $|t| \leq t_0$ whereas this is not a factor in bounding $Op_{\gamma_1 \dots \gamma_M; y}(\rho_q^{\text{new}})$.

where $C > 0$ is a constant and

$$\mu_1 = \sum_{\gamma_{M+1}, \dots, \gamma_N} e^{r_\sigma^{(\tilde{M})}(\alpha_{\gamma_{M+1} \dots \gamma_N o})} h_{\gamma_N} h_{\gamma_{N-1}} \dots h_{\gamma_{M+1}}. \quad (4.8)$$

Note that μ_1 is a real measure and we can estimate its ℓ^1 norm

$$\|\mu_1\|_1 \leq \sum_{\gamma_{M+1}, \dots, \gamma_N} e^{r_\sigma^{(\tilde{M})}(\alpha_{\gamma_{M+1} \dots \gamma_N o})} = \mathcal{L}_\sigma^{\tilde{M}}[1] \ll \lambda_\sigma^{\tilde{M}}. \quad (4.9)$$

Let $\tilde{\mu}(g) := \overline{\mu(g^{-1})}$. We will obtain a bound for $\mu_{\gamma_1 \dots \gamma_M; y}$ by getting a bound for $\tilde{\mu}_1 * \mu_1$ and then using the following lemma, the analogous version of [20, Lemma 44].

Lemma 4.6. *Given (4.7), we have for some $D > 0$ a constant depending only on g*

$$\|\mu_{\gamma_1 \dots \gamma_M; y}\|_{\ell_{\text{new}}^2(\Gamma_q)} \leq C' e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M o})} \left(\frac{|\Gamma_q| \|\tilde{\mu}_1 * \mu_1\|_2^2}{q^D} \right)^{\frac{1}{4}}.$$

Here $\|\tilde{\mu}_1 * \mu_1\|_2^2$ denotes the ℓ^2 norm of the measure on Γ_q and $\|\mu_{\gamma_1 \dots \gamma_M; y}\|_{\ell_{\text{new}}^2(\Gamma_q)}$ is the operator norm of $\mu_{\gamma_1 \dots \gamma_M; y}$ acting on the new subspace of $\ell^2(\Gamma_q)$.

Proof. We need to use the lower bound for the degree of new irreducible representations of $\text{Sp}((\mathbf{Z}/q\mathbf{Z})^{2g}, \omega_\pi)$ that is given in Proposition 1.9. Supposing that the smallest new irreducible representation has dimension $\gg q^D$ then by the trace formula argument of [20, Lemma 44] the largest eigenvalue of A^*A where $A := \mu_{\gamma_1 \dots \gamma_M; y}^*$ acting on $\ell_{\text{new}}^2(\Gamma_q)$ satisfies

$$\lambda^2 q^D \leq C' (e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M o})})^4 |\Gamma_q| \|\tilde{\mu}_1 * \mu_1\|_2^2.$$

The crucial point is that the eigenvalue appears with high multiplicity in the trace formula, an idea that goes back to Sarnak and Xue [28]. Since $\|A\| = \lambda^{1/2}$ the lemma follows. \square

We will also use the following lemma.

Lemma 4.7 ([20, Proposition 45]). *Given that $\|\mu_1 * \phi\|_{\ell^2(\Gamma_q)} \leq (1 - \epsilon_1)^{\tilde{M}} \|\phi\|_{\ell^2(\Gamma_q)}$ for some positive ϵ_1 , for all $\tilde{M} > 0$ and for all $\phi \in \ell_0^2(\Gamma_q)$, we have for some $c > 0$ that if $\tilde{M} \geq c \log q$*

$$\|\tilde{\mu}_1 * \mu_1\|_2 \leq 2 \frac{\|\mu_1\|_1^2}{|\Gamma_q|^{1/2}}.$$

The chain of bounds we will use to prove Proposition 4.5 are

$$\begin{aligned} \|\mu_1 * \phi\| &\leq (1 - \epsilon_1)^{\tilde{M}} \|\phi\|, \phi \in \ell_0^2(\Gamma_q) \\ &\stackrel{\substack{\text{Lemma 4.7} \\ \tilde{M} \approx c \log q}}{\rightsquigarrow} \|\tilde{\mu}_1 * \mu_1\|_2 \leq 2 \frac{\|\mu_1\|_1^2}{|\Gamma_q|^{1/2}} \\ &\stackrel{\substack{\text{Lemma 4.6} \\ (4.9)}}{\rightsquigarrow} \|\mu_{\gamma_1 \dots \gamma_M; y}\|_{\ell_{\text{new}}^2(\Gamma_q)} \leq C'' e^{r_\sigma^{(M)}(\alpha_{\gamma_1 \dots \gamma_M o})} q^{-D/4} \|\mu_1\|_1 \\ &\qquad\qquad\qquad \text{Proposition 4.5.} \end{aligned}$$

It remains to prove the first inequality above, this is the topic of the next section.

4.4 Decoupling II: Bounding the real measure μ_1 .

Lemma 4.8. *There is some $\epsilon_1 > 0$ such that for all $\phi \in \ell_0^2(\Gamma_q)$ we have*

$$\|\mu_1 * \phi\|_{\ell^2} \leq (1 - \epsilon_1)^{\tilde{M}} \|\phi\|_{\ell^2}.$$

To prove Lemma 4.8 we adapt the arguments of [20, Appendix] to the infinite alphabet setting, using also different spectral gap inputs (Theorems 1.7 and 1.8) that rely on our preparation of the set Υ and its relation to the Zariski-dense Rauzy-Veech group G_π .

We further decompose

$$\tilde{M} = LR \tag{4.10}$$

where L is going to be chosen to be a large constant, and decompose $\{M+1, \dots, N\}$ into blocks of size either 1, $L-1$ or L . Let

$$I_{i,j} = [\gamma_i, \gamma_{i+j}]$$

denote the block of all $\gamma_{i'}$ with $i \leq i' \leq j$. Rewrite the summation in (4.8) as

$$\mu_1 = \sum_{I_{M+1, M+L-1}, I_{M+L+1, M+2L-1}, \dots, I_{N-L+1, N-1}} \sum_{\gamma_{M+L}, \gamma_{M+2L}, \dots, \gamma_N} e^{r_\sigma^{(\tilde{M})}(\alpha_{\gamma_{M+1} \dots \gamma_N} o)} h_{\gamma_N} h_{\gamma_{N-1}} \dots h_{\gamma_{M+1}}. \tag{4.11}$$

This reordering of summation is permitted since the sums are suitably absolutely convergent by Theorem 4.4 and the following discussion. Following [20, (A.15)], using contraction properties of α_{γ_i} and the bound (4.2) for the derivative of $r_\sigma^{(\tilde{M})}$, one has the bound

$$e^{r_\sigma^{(\tilde{M})}(\alpha_{\gamma_{M+1} \dots \gamma_N} o)} \leq \exp(c\Lambda^{-L})^{R-1} \beta_1 \beta_2 \dots \beta_R \tag{4.12}$$

where

$$\beta_R = e^{r_\sigma^{(L)}(\alpha_{\gamma_{N-L+1} \dots \gamma_N} o)}, \quad \beta_j = e^{r_\sigma^{(L)}(\alpha_{\gamma_{M+(j-1)L+1} \dots \gamma_{M+(j+1)L-1} o)}}, \quad 1 \leq j \leq R-1.$$

and $c > 0$ is a constant. Notice the important feature that each β_j depends on only one of γ_{M+jL} . Inserting (4.12) into (4.11) gives

$$\mu_1 \leq \exp(c\Lambda^{-L})^{R-1} \sum_{I_{M+1, M+L-1}, I_{M+L+1, M+2L-1}, \dots, I_{N-L+1, N-1}} \eta_R * \eta_{R-1} * \dots * \eta_1 \tag{4.13}$$

where the $\eta_j = \eta_j(I_{M+1, M+L-1}, I_{M+L+1, M+2L-1}, \dots, I_{N-L+1, N-1})$ are given by

$$\begin{aligned} \eta_R &:= \sum_{\gamma_N} \beta_R(\gamma_{N-L}, \dots, \gamma_N) h_{\gamma_N} \dots h_{\gamma_{N-L+1}}, \\ \eta_j &:= \sum_{\gamma_{M+jL}} \beta_j(\gamma_{M+(j-1)L+1}, \dots, \gamma_{M+(j+1)L-1}) h_{\gamma_{M+jL}} \dots h_{\gamma_{M+(j-1)L+1}}, \quad 1 \leq j \leq R-1. \end{aligned}$$

We now aim for bounds on the operator norms of the measures η_j acting by convolution on $\ell_0^2(\Gamma_q)$. We write $\|\eta_j\|_{op}$ for this operator norm. Consider, taking for example $1 \leq j \leq R-1$

$$\eta_j * \tilde{\eta}_j = \sum_{\gamma_{M+jL}, \gamma'_{M+jL}} \beta_j(\dots, \gamma_{M+jL}, \dots) \beta_j(\dots, \gamma'_{M+jL}, \dots) h_{\gamma_{M+jL}} (h_{\gamma'_{M+jL}})^{-1}. \quad (4.14)$$

Since

$$\|\eta_j\|_{op} = \|\tilde{\eta}_j\|_{op} = \sup_{\phi \in \ell_0^2(\Gamma_q): \|\phi\|=1} \langle \eta_j * \tilde{\eta}_j \phi, \phi \rangle^{1/2} \quad (4.15)$$

we turn to estimating the operator norm of $\eta_j * \tilde{\eta}_j$ on $\ell_0^2(\Gamma_q)$. We need to both

1. estimate the values of β_j and
2. discuss the group elements $h_{\gamma_{M+jL}} (h_{\gamma'_{M+jL}})^{-1}$.

These are both points of departure from [20, Appendix], so we give more details.

1) Continuing with $1 \leq j \leq R-1$ (the edge case $j = R$ is similar) we have

$$\begin{aligned} \beta_j(\dots, \gamma_{M+jL}, \dots) &= e^{r_\sigma^{(L)}(\alpha_{\gamma_{M+(j-1)L+1} \dots \gamma_{M+(j+1)L-1}} o)} \\ &= \exp \left(\sum_{i=0}^{L-2} r_\sigma(\alpha_{\gamma_{M+(j-1)L+1+i} \dots \gamma_{M+(j+1)L-1}} o) \right) \exp \left(r_\sigma(\alpha_{\gamma_{M+jL} \dots \gamma_{M+(j+1)L-1}} o) \right) \\ &= \exp \left(\sum_{i=0}^{L-2} r_\sigma(\alpha_{\gamma_{M+(j-1)L+1+i} \dots \gamma_{M+jL-1}} o) + O(\Lambda^{-i}) \right) \exp \left(r_\sigma(\alpha_{\gamma_{M+jL} \dots \gamma_{M+(j+1)L-1}} o) \right) \\ &\asymp B(\gamma_{M+(j-1)L+1+i}, \dots, \gamma_{M+jL-1}) \exp \left(r_\sigma(\alpha_{\gamma_{M+jL} \dots \gamma_{M+(j+1)L-1}} o) \right) \end{aligned} \quad (4.16)$$

where \asymp means bounded above and below by a constant independent of all γ_i and L , and

$$B(\gamma_{M+(j-1)L+1+i}, \dots, \gamma_{M+jL-1}) := \exp \left(r_\sigma^{(L-1)}(\alpha_{\gamma_{M+(j-1)L+1+i} \dots \gamma_{M+jL-1}} o) \right).$$

Note the arguments of B are fixed given η_j . Also note that for fixed η_j the values

$$\alpha_{\gamma_{M+jL+1} \dots \gamma_{M+(j+1)L-1}} o$$

lie in a (cylinder) set $U(\eta_j)$ with diameter $\ll \Lambda^{-L}$. On the other hand, the derivative of $r_\sigma \circ \alpha_{\gamma_{M+jL}}$ is uniformly bounded so the values in the exponent of (4.16) fluctuate by at most $\ll \Lambda^{-L}$ while $\alpha_{\gamma_{M+jL}}$ is fixed. Therefore

$$\beta_j(\dots, \gamma_{M+jL}, \dots) \asymp B(\gamma_{M+(j-1)L+1+i}, \dots, \gamma_{M+jL-1}) \exp \left(r_\sigma(\alpha_{\gamma_{M+jL}} o) \right). \quad (4.17)$$

In light of this estimate and the discussion after Theorem 4.4 concerning convergence of infinite sums, we see that η_j and $\eta_j * \tilde{\eta}_j$ have finite ℓ_1 norms. This supports our earlier justification of reordering of summations.

2) Recall Υ from Lemma 4.1. We can write

$$\eta_j * \tilde{\eta}_j = \nu + \tilde{\nu}$$

where ν is the contribution to (4.14) from $\gamma_{M+jL}, \gamma'_{M+jL} \in \gamma_0 \cdot \Upsilon$ and $\tilde{\nu}$ are the remaining contributions. Then the support of ν is the reduction mod q of the set

$$\Sigma = \{\Theta_{\gamma}^* (\Theta_{\gamma'}^*)^{-1} : \gamma, \gamma' \in \gamma_0 \cdot \Upsilon\}.$$

By Lemma 4.1, the set Σ generates the conjugate of G'_π by $\Theta_{\gamma_0}^*$. Call this conjugate group G''_π .

We now bring these arguments 1) and 2) together. Let $\nu = \eta_j * \tilde{\eta}_j$. Note that the operator formed from convolution by ν on $\ell_0^2(\Gamma_q)$ is self-adjoint and positive. Therefore the operator norm of $\|\nu\|$ acting by convolution on $\ell_0^2(\Gamma_q)$ is bounded by

$$\|\nu\|_{op} \leq \sup_{\phi \in \ell_0^2(\Gamma_q), \|\phi\|=1} \langle \nu * \phi, \phi \rangle. \quad (4.18)$$

We need to use the following property of the action of G'_π on Γ_q .

Lemma 4.9 (No almost invariant vectors). *There is finite q_0 and some $\epsilon > 0$ such that for all q such that*

- *q is coprime to q_0 if G_π is finite index in $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_\pi)$,*
- *q is a squarefree positive integer coprime to q_0 if G_π is Zariski-dense but not finite index,*

the following holds. For all $\phi \in \ell_0^2(\Gamma_q)$ with $\|\phi\|_{\ell^2} = 1$ there is some $g \in \Sigma$ such that if $g_q := g \bmod q$ then

$$\|g_q * \phi - \phi\|_{\ell^2} > \epsilon.$$

Proof. Choose q_0 so that strong approximation holds for G''_π at all q coprime to q_0 . If G_π and hence G''_π is finite index in $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_\pi)$ then the lemma is a direct result of Kazhdan's property (T) for the group $\mathrm{Sp}(\mathbf{Z}^{2g}, \omega_\pi)$ (Theorem 1.7). In the case G''_π is only Zariski-dense then we rely instead on the result of Salehi Golsefidy and Varjú from Theorem 1.8. Indeed, a standard argument translates Theorem 1.8 into the non-existence of invariant vectors statement of the lemma. \square

Write

$$\nu = \sum_{g_q \in \Gamma_q} \nu_{g_q} g_q.$$

Let ϵ, g_q^0 be the constant (resp. group element) provided by Lemma 4.9 on inputting ϕ with $\|\phi\| = 1$. Then it is straightforward to check that $|\Re(\langle g_q^0 * \phi, \phi \rangle)| < (1 - \epsilon')$ where $\epsilon' = \epsilon^2/2$. Returning to (4.18), using $\nu_{g_q} = \nu_{(g_q)^{-1}}$ from (4.14) we get

$$\begin{aligned} \langle \nu * \phi, \phi \rangle &= \sum_{g_q \in \Gamma_q} \nu_{g_q} \langle g_q * \phi, \phi \rangle = \sum_{g_q \in \Gamma_q} \nu_{g_q} \Re \langle g_q * \phi, \phi \rangle \\ &= \nu_{g_q^0} \Re \langle g_q^0 * \phi, \phi \rangle + \sum_{g_q \neq g_q^0} \nu_{g_q} \Re \langle g_q * \phi, \phi \rangle \\ &\leq (1 - \epsilon') \nu_{g_q^0} + \sum_{g_q \neq g_q^0} \nu_{g_q} = \|\nu\|_1 - \epsilon' \nu_{g_q^0}. \end{aligned}$$

Also, from (4.17), and the convergence of $\sum_{\gamma_{M+jL}} \exp(r_\sigma(\alpha_{\gamma_{M+jL}} o))$ we get

$$\nu_{g_q^0} \geq C \|\nu\|_1$$

with constant C independent of $\nu_{g_q^0}$ and η_j . So combining this with the preceding estimate and (4.18) we get

$$\|\nu\|_{op} \leq \|\nu\|_1 (1 - \epsilon'')$$

for some $\epsilon'' > 0$. Inserting this into (4.15) gives

$$\|\eta_j\|_{op} \leq \|\eta_j\|_{\ell^1} (1 - \epsilon'')^{1/2}. \quad (4.19)$$

Using (4.19) in (4.13) gives for any $\phi \in \ell_0^2(\Gamma_q)$

$$\|\mu_1 * \phi\|_{\ell^2} \leq \exp(c\Lambda^{-L})^{R-1} (1 - \epsilon'')^{R/2} \|\phi\|$$

and we now choose L so the constant above is $\leq (1 - \epsilon)^R$ for some $\epsilon > 0$. Recalling (4.10), this completes the proof of Lemma 4.8. Therefore using the chain of bounds at the bottom of Section 4.3 we have proved Proposition 4.5.

5 Quasirandomness

In this section we show how one may obtain Proposition 1.9. We follow the type of argument given by Kelmer and Silberman in [15, Section 4] for rank one groups (see also [21] for a small improvement to that argument). We may treat the group $\mathrm{Sp}_{2g}(\mathbf{Z})$ without loss of generality, that is, we assume the symplectic form is the standard one. Let $g \geq 2$. Let $q \in \mathbf{N}$ and let (ρ, V) be an irreducible unitary representation of $\mathrm{Sp}_{2g}(\mathbf{Z}/q\mathbf{Z})$ that is not obtained by a composition

$$\mathrm{Sp}_{2g}(\mathbf{Z}/q\mathbf{Z}) \rightarrow \mathrm{Sp}_{2g}(\mathbf{Z}/q_1\mathbf{Z}) \xrightarrow{\rho'} U(V)$$

with $q_1|q$. We refer to this property as ρ being *new*.

5.1 The case when q is prime

For p an odd prime, let \mathbb{F}_p denote the finite field with p elements. The table of Seitz and Zalesskii in [29, Table 1] implies that $P\mathrm{Sp}_{2g}(\mathbb{F}_p)$ has no projective complex irreducible representation of dimension $< \frac{1}{2}(p^g - 1)$ and hence this is also a lower bound for the dimension of an irreducible representation of $\mathrm{Sp}_{2g}(\mathbb{F}_p)$.

5.2 The case $q = p^r$

In this section we prove the following

Proposition 5.1. *There is some $C > 0$ depending only on g such that for all $r \geq 2$, letting $R := \lfloor r/2 \rfloor$ any new representation (ρ, V) of $\mathrm{Sp}_{2g}(\mathbf{Z}/p^r\mathbf{Z})$ has dimension at least*

$$\dim \rho \geq Cp^R.$$

Let $q = p^r$. Write $H_q := \mathrm{Sp}_{2g}(\mathbf{Z}/q\mathbf{Z})$ and for $q'|q$ let $H_q(q')$ be the kernel of the reduction modulo q' map

$$H_q \rightarrow H_{q'}.$$

Let $\mathfrak{g}(\mathbf{Z}/q\mathbf{Z})$ denote the Lie algebra of Sp_{2g} over $\mathbf{Z}/q\mathbf{Z}$. We view this as an abelian group. Let $R = \lfloor r/2 \rfloor$. The congruence subgroup $H_{p^r}(p^{r-R})$ is an abelian normal subgroup of H_{p^r} that is naturally isomorphic to $\mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$. The action of H_{p^r} on $H_{p^r}(p^{r-R})$ by conjugation descends to an action of H_{p^R} . After using the isomorphism $H_{p^r}(p^{r-R}) \cong \mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$ this conjugation action is identified with the Adjoint action of H_{p^R} on $\mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$, i.e.

$$\mathrm{Ad}(g)v = gvg^{-1}, \quad g \in H_{p^R}, v \in \mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z}).$$

Let (ρ, V) be a unitary representation of H_q . Suppose $R \geq 1$. If ρ is trivial when restricted to $H_q(p^R)$ then ρ is not a new representation. More generally, if ρ is new, then the restriction of ρ to $H_q(p^{r-R})$ must not be trivial on any $H_q(p^{r-R+\eta})$ with $\eta \in \mathbf{Z}_+$ since these are also normal subgroups with $H_q/H_q(p^{r-R+\eta}) \cong H_{p^{r-R+\eta}}$. Notice $H_q(p^{r-R+\eta}) \leq H_q(p^{r-R})$ corresponds to the inclusion $p^\eta \mathfrak{g}(\mathbf{Z}/p^{R-\eta}\mathbf{Z}) \leq \mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$.

The strategy is to consider the H_{p^R} invariant set of characters of $\mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$ that appear when restricting ρ to $H_{p^r}(p^{r-R}) \cong \mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$, since the size of this set gives a lower bound for the dimension of ρ .

The Killing form on $\mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$ is non-degenerate which allows us to identify the unitary dual $\widehat{\mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})}$ with $\mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$. Under this identification, the co-Adjoint action on characters becomes an Adjoint action on $\mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$. Moreover any character that is non trivial on each $H_q(p^{r-R+\eta})$, $\eta \in \mathbf{Z}_+$, becomes an element of $\mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$ which is not $\equiv 0 \pmod p$.

We have therefore reduced Proposition 5.1 therefore to the following Lemma.

Lemma 5.2. *There is some $C > 0$ depending only on g such that for all $R \geq 1$ the H_{p^R} -Adjoint orbit of any $X \in \mathfrak{g}(\mathbf{Z}/p^R\mathbf{Z})$ with $X \not\equiv 0 \pmod p$ has size*

$$|\mathrm{Ad}(H_{p^R}).X| \geq Cp^R.$$

Proof. By orbit-stabilizer theorem the orbit has size at least

$$\frac{|H_{p^R}|}{|C_{H_{p^R}}(X)|} \tag{5.1}$$

where we write C to stand for centralizer, therefore $C_{H_{p^r}}(X) = \{h \in H_{p^r} : hXh^{-1} = X\}$. Since H_{p^R} is an $R-1$ fold extension of H_p by groups isomorphic to $\mathfrak{g}(\mathbb{F}_p)$ we know $|H_{p^R}| = |H_p||\mathfrak{g}(\mathbb{F}_p)|^{R-1} \gg p^{R \cdot \dim(\mathrm{Sp}_{2g})} = p^{g(2g+1)R}$. This gives the bound we will use for the numerator of (5.1).

Considering next the denominator of (5.1), by an elementary induction argument appearing in [15, Proof of Proposition 4.3]

$$|C_{H_{p^r}}(X)| \leq |C_{H_p}(X \bmod p)| |C_{\mathfrak{g}(\mathbb{F}_p)}(X \bmod p)|^{R-1} \tag{5.2}$$

where the latter centralizer is $C_{\mathfrak{g}(\mathbb{F}_p)}(X \bmod p) = \{y \in \mathfrak{g}(\mathbb{F}_p) : [y, X \bmod p] \equiv 0\}$. According to Springer and Steinberg [31, II. 4.1, 4.2, IV. 2.26], the algebraic group $C_{\mathrm{Sp}_{2g}}(X \bmod p)$

defined over \mathbb{F}_p has a number of components bounded by a constant depending only on g . By a bound of Nori [25, Lemma 3.5] each component can have at most $\leq (p+1)^{\dim C_{\mathrm{Sp}_{2g}}(X \bmod p)}$ points over \mathbb{F}_p . But $\dim C_{\mathrm{Sp}_{2g}}(X \bmod p)$ is also the dimension of the centralizer of $X \bmod p$ in $\mathfrak{g}(\mathbb{F}_p)$ so we have now reduced the estimation of the right hand side of (5.2) to a bound for

$$\dim C_{\mathfrak{g}(\mathbb{F}_p)}(X')$$

where $X' = X \bmod p$ is a nonzero element of $\mathfrak{g}(\mathbb{F}_p)$.

Assume the bound $\dim C_{\mathfrak{g}(\mathbb{F}_p)}(X') \leq \dim(\mathrm{Sp}_{2g}) - e = g(2g+1) - e$. Then putting our previous estimates together the orbit has size at least

$$\gg \frac{p^{g(2g+1)R}}{(p+1)^{g(2g+1)-e} (p^{g(2g+1)-e})^{R-1}} \gg p^{eR}.$$

Since it is not particularly important here to optimize e , we give the easy argument that one may take $e = 1$ since $\mathfrak{g}(\mathbb{F}_p)$ has no nontrivial center¹⁵. This gives the result stated in the lemma. □

5.3 The case of general moduli

If p_i are primes and

$$q = \prod_{i=1}^M p_i^{m_i}$$

is the prime factorization of q , then we have by the Chinese remainder theorem

$$\mathrm{Sp}_{2g}(\mathbf{Z}/q\mathbf{Z}) \cong \prod_{i=1}^M \mathrm{Sp}_{2g}(\mathbf{Z}/p_i^{m_i}\mathbf{Z}).$$

Then ρ splits as a tensor product

$$\rho = \bigotimes_{i=1}^M \rho_i$$

where ρ_i are irreducible representations of $\mathrm{Sp}_{2g}(\mathbf{Z}/p_i^{m_i}\mathbf{Z})$. Since ρ is new, all of the ρ_i are new. Now using Proposition 5.1 and the bounds for the case of prime modulus from Section 5.1 gives

$$\dim \rho \geq \prod_{i: m_i=1} \frac{1}{2}(p_i^g - 1) \prod_{i: m_i>1} C p_i^{\lfloor m_i/2 \rfloor} \geq q^{1/2} (C')^{-\omega(q)}$$

given $g \geq 2$ for some $C' > 1$ and $\omega(q)$ standing for the number of distinct prime factors of q . But $(C')^{\omega(q)} \ll_{\epsilon} q^{\epsilon}$ for any $\epsilon > 0$. This concludes the proof of Proposition 1.9, in fact, our proof shows that one may take D as close as one likes to $1/2$ provided one chooses C appropriately.

¹⁵One may definitely do better here, and it would be good to work out the best possible bound, but it is not the purpose of the current paper.

References

- [1] A. Avila, C. Matheus, and J.-C. Yoccoz. Zorich conjecture for hyperelliptic Rauzy-Veech groups. *arXiv:1606.01227*, June 2016.
- [2] Artur Avila and Sébastien Gouëzel. Small eigenvalues of the Laplacian for algebraic measures in moduli space, and mixing properties of the Teichmüller flow. *Ann. of Math. (2)*, 178(2):385–442, 2013.
- [3] Artur Avila, Sébastien Gouëzel, and Jean-Christophe Yoccoz. Exponential mixing for the Teichmüller flow. *Publ. Math. Inst. Hautes Études Sci.*, (104):143–211, 2006.
- [4] Artur Avila and Marcelo Viana. Simplicity of Lyapunov spectra: proof of the Zorich-Kontsevich conjecture. *Acta Math.*, 198(1):1–56, 2007.
- [5] V. Bargmann. Irreducible unitary representations of the Lorentz group. *Ann. of Math. (2)*, 48:568–640, 1947.
- [6] Y. Benoist. Propriétés asymptotiques des groupes linéaires. *Geom. Funct. Anal.*, 7(1):1–47, 1997.
- [7] Armand Borel. Density properties for certain subgroups of semi-simple groups without compact components. *Ann. of Math. (2)*, 72:179–188, 1960.
- [8] Jean Bourgain and Alex Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Ann. of Math. (2)*, 167(2):625–642, 2008.
- [9] Jean Bourgain, Alex Gamburd, and Peter Sarnak. Generalization of Selberg’s $\frac{3}{16}$ theorem and affine sieve. *Acta Math.*, 207(2):255–290, 2011.
- [10] Dmitry Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math. (2)*, 147(2):357–390, 1998.
- [11] S. Filip. Zero Lyapunov exponents and monodromy of the Kontsevich-Zorich cocycle. *arXiv:1410.2129*, October 2014.
- [12] A. Salehi Golsefidy and Péter P. Varjú. Expansion in perfect groups. *Geom. Funct. Anal.*, 22(6):1832–1891, 2012.
- [13] W. T. Gowers. Quasirandom groups. *Combin. Probab. Comput.*, 17(3):363–387, 2008.
- [14] D. A. Každan. On the connection of the dual space of a group with the structure of its closed subgroups. *Funkcional. Anal. i Priložen.*, 1:71–74, 1967.
- [15] Dubi Kelmer and Lior Silberman. A uniform spectral gap for congruence covers of a hyperbolic manifold. *Amer. J. Math.*, 135(4):1067–1085, 2013.
- [16] Henry H. Kim. Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2 . *J. Amer. Math. Soc.*, 16(1):139–183, 2003. With appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak.

- [17] Anthony W. Knap. *Representation theory of semisimple groups*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2001. An overview based on examples, Reprint of the 1986 original.
- [18] Maxim Kontsevich and Anton Zorich. Connected components of the moduli spaces of Abelian differentials with prescribed singularities. *Invent. Math.*, 153(3):631–678, 2003.
- [19] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.
- [20] M. Magee, H. Oh, and D. Winter. Uniform congruence counting for Schottky semi-groups in $\mathrm{SL}_2(\mathbf{Z})$, with Appendix by J. Bourgain, A. Kontorovich and M. Magee. *arXiv:1601.03705*, January 2016.
- [21] Michael Magee. Quantitative spectral gap for thin groups of hyperbolic isometries. *J. Eur. Math. Soc. (JEMS)*, 17(1):151–187, 2015.
- [22] S. Marmi, P. Moussa, and J.-C. Yoccoz. The cohomological equation for Roth-type interval exchange maps. *J. Amer. Math. Soc.*, 18(4):823–872 (electronic), 2005.
- [23] Howard Masur. Interval exchange transformations and measured foliations. *Ann. of Math. (2)*, 115(1):169–200, 1982.
- [24] C. R. Matthews, L. N. Vaserstein, and B. Weisfeiler. Congruence properties of Zariski-dense subgroups. I. *Proc. London Math. Soc. (3)*, 48(3):514–532, 1984.
- [25] Madhav V. Nori. On subgroups of $\mathrm{GL}_n(\mathbf{F}_p)$. *Invent. Math.*, 88(2):257–275, 1987.
- [26] Hee Oh and Dale Winter. Uniform exponential mixing and resonance free regions for convex cocompact congruence subgroups of $\mathrm{SL}_2(\mathbf{Z})$. *J. Amer. Math. Soc.*, 29(4):1069–1115, 2016.
- [27] Gérard Rauzy. Échanges d’intervalles et transformations induites. *Acta Arith.*, 34(4):315–328, 1979.
- [28] Peter Sarnak and Xiao Xi Xue. Bounds for multiplicities of automorphic representations. *Duke Math. J.*, 64(1):207–227, 1991.
- [29] Gary M. Seitz and Alexander E. Zalesskii. On the minimal degrees of projective representations of the finite Chevalley groups. II. *J. Algebra*, 158(1):233–243, 1993.
- [30] Atle Selberg. On the estimation of Fourier coefficients of modular forms. In *Proc. Sympos. Pure Math., Vol. VIII*, pages 1–15. Amer. Math. Soc., Providence, R.I., 1965.
- [31] T. A. Springer and R. Steinberg. Conjugacy classes. In *Seminar on Algebraic Groups and Related Finite Groups (The Institute for Advanced Study, Princeton, N.J., 1968/69)*, Lecture Notes in Mathematics, Vol. 131, pages 167–266. Springer, Berlin, 1970.
- [32] William A. Veech. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math. (2)*, 115(1):201–242, 1982.

- [33] M. Viana. Dynamics of Interval Exchange Transformations and Teichmüller Flows.
<http://w3.impa.br/~viana/out/ietf.pdf>.
- [34] Anton Zorich. Finite Gauss measure on the space of interval exchange transformations. Lyapunov exponents. *Ann. Inst. Fourier (Grenoble)*, 46(2):325–370, 1996.
- [35] Anton Zorich. How do the leaves of a closed 1-form wind around a surface? In *Pseudoperiodic topology*, volume 197 of *Amer. Math. Soc. Transl. Ser. 2*, pages 135–178. Amer. Math. Soc., Providence, RI, 1999.
- [36] Anton Zorich. Flat surfaces. In *Frontiers in number theory, physics, and geometry. I*, pages 437–583. Springer, Berlin, 2006.

Michael Magee,
 Department of Mathematics,
 Yale University,
 PO Box 208283, New Haven, CT 06520 USA
michael.magee@yale.edu